PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 :
A61K 35/14, C12P 21/00, 21/02
C12N 15/00

A1

(11) International Publication Number: WO 88/ 00831

(43) International Publication Date: 11 February 1988 (11.02.88)

(21) International Application Number: PCT/US87/01814

(22) International Filing Date: 31 July 1987 (31.07.87)

(31) Priority Application Number: 893,375

(32) Priority Date: 1 August 1986 (01.08.86)

(33) Priority Country: US

(60) Parent Application or Grant
(63) Related by Continuation
US 893,375 (CIP)
Filed on 1 August 1986 (01.08.86)

(71) Applicant *(for all designated States except US)*: BIOGEN N.V. [NL/NL]; Pietermaai 15, Willemstad, Curaçao (AN).

(71)(72) Applicant and Inventor: PASEK, Mark, P. [US/US]; 177 Lexington Street, Belmont, MA 02178 (US).

(74) Agents: HALEY, James, F., Jr. et al.; Fish & Neave, 875 Third Avenue, New York, NY 10022-6250 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, KR, LU (European patent), NL (European patent), SE (European patent), US.

Published
With international search report.

(54) Title: DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE POLYPEPTIDES IN HIGH YIELDS

(57) Abstract

DNA sequences coding on expression for modified factor VIII:C and modified factor VIII:C-like polypeptides and methods of making them in high yields in appropriate hosts transformed with those DNA sequences. DNA sequences containing internal deletions removing a major part of the sequence which codes on expression for the maturation polypeptide of factor VIII:C express modified factor VIII:C and modified factor VIII:C-like polypeptides 20 times more efficiently than DNA sequences coding for the factor VIII:C.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FR**	France	**ML**	Mali
AU	Australia	**GA**	Gabon	**MR**	Mauritania
BB	Barbados	**GB**	United Kingdom	**MW**	Malawi
BE	Belgium	**HU**	Hungary	**NL**	Netherlands
BG	Bulgaria	**IT**	Italy	**NO**	Norway
BJ	Benin	**JP**	Japan	**RO**	Romania
BR	Brazil	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SN**	Senegal
CH	Switzerland	**LI**	Liechtenstein	**SU**	Soviet Union
CM	Cameroon	**LK**	Sri Lanka	**TD**	Chad
DE	Germany, Federal Republic of	**LU**	Luxembourg	**TG**	Togo
DK	Denmark	**MC**	Monaco	**US**	United States of America
FI	Finland	**MG**	Madagascar		

DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE POLYPEPTIDES IN HIGH YIELDS

TECHNICAL FIELD OF THE INVENTION

This invention relates to DNA sequences coding for modified factor VIII:C-like polypeptides and processes for producing them using those DNA sequences. More particularly, this invention relates to the production of modified factor VIII:C and modified factor VIII:C-like polypeptides which display the biological activity of factor VIII:C. In addition, the polypeptides of this invention are produced in higher yields than previously produced factor VIII:C-like polypeptides and are more easily purified into biochemically pure mature factor VIII:C.

BACKGROUND OF THE INVENTION

Factor VIII:C, a large plasma glycoprotein, functions as the procoagulant component of factor VIII, which plays an integral role in the cascade mechanism of blood coagulation [see generally, W. J. Williams et al., Hematology, pp. 1085-90, McGraw-Hill, New York (1972)]. Factor VIII:C circulates in the blood as a complex with factor VIIIR:Ag (also known as von Willebrand factor protein) which is a large

protein associated with platelet aggregation and adhesive properties.

Factor VIII:C is synthesized as a single chain macromolecular precursor, which is later cleaved to yield the fragments which constitute "mature" factor VIII:C. Mature factor VIII:C is composed of two chains bridged by a calcium ion; an amino-terminal heavy chain of 740 amino acids, and a carboxy-terminal light chain of 684 amino acids. The primary translation product of factor VIII:C is a single chain in which the heavy chain of mature factor VIII:C is separated from the light chain by a "maturation polypeptide" of 908 amino acids. The excision of this maturation polypeptide is initiated by proteolytic cleavage of the primary translation product by an unknown or yet unidentified protease at the Arg 1648 - Glu 1649 peptide bond. The initial nick event begins a series of successive proteolytic cleavages which shorten the nascent heavy chain from its carboxy terminus. Eventually the mature heavy chain of 740 amino acids results and in combination with the light chain of 684 amino acids, comprises mature factor VIII:C [see L.-O. Andersson et al. "Isolation and Characterization of Human Factor VIII: Molecular Forms In Commercial Factor VIII Concentrate, Cryoprecipitate, and Plasma," PNAS(USA), 83, pp. 2979-83 (1986)]. This complex is then activated by thrombin by cleavage at the Arg 1689-Ser 1690 bond [D. Eaton et al., Biochemistry, 25, pp. 505-12 (1986)].

Haemophilia A is a sex-linked hemorrhagic disease which is caused by a deficiency, either in amount or in biological activity, of factor VIII:C. The symptoms of acutely bleeding haemophilia patients are treated with factor VIII traditionally purified from normal sera. Various methods of purification have been described in the literature [see, Zimmerman

et al., United States patent 4,361,509; Saundrey et al. United States patent 4,578,218; E.G.D. Tuddenhem et al., "The Properties of Factor VIII Coagulant Activity Prepared By Immunoadsorbent Chromatography, Journal of Laboratory Clinical Medicine, 93, pp. 40-53 (1979); D. E. G. Austen, "The Chromatographic Separation of Factor VIII on Aminohexyl Sepharose," British Journal of Hematology, 43, pp. 669-74 (1979); M. Weinstein et al., "Analysis of Factor VIII Coagulant Antigen In Normal, Thrombin-treated, and Hemophilic Plasma," PNAS (USA), 78, pp. 5137-41 (1981); P. J. Fay et al., "Purification And Characterization Of A Highly Purified Human Factor VIII Consisting Of A Single Type Of Polypeptide Chain," PNAS (USA), 79, pp. 7200-04 (1982); C. A. Fulcher and T. S. Zimmerman, "Characterization Of The Human Factor VIII Procoagulant Protein With A Heterologous Precipitating Antibody," PNAS (USA), 79, pp. 1648-52 (1982); F. Rotblat et al., Thromb. Haemostasis, 50, p. 108 (1983); C. A. Fulcher et al., Blood, 61, pp. 807-11 (1983)].

However, purification has proven to be difficult because of the relatively low concentration of factor VIII:C in serum, its tight association with the larger factor VIIIR:Ag and its sensitivity to degradation by serum proteases. Factor VIII:C when purified from plasma thus contains a heterogeneous mixture of heavy chains ranging in length from 1648 amino acids down to 740 amino acids which result from these numerous proteolytic events [Andersson et al., supra, p. 2983]. The heterogenous mixture of chains observed in plasma-purified factor VIII:C, has made recovery of a substantially pure mature factor VIII:C almost impossible. Furthermore, traditional treatment of haemophilia with factor VIII purified from plasma has serious drawbacks. Specifically, it can lead to the unintended transfer of the

causative agents of hepatitis or the virus associated with Acquired Immune Deficiency Syndrome.

In view of its importance in the treatment of haemophilia, numerous attempts have been made to produce large quantities of factor VIII:C using recombinant DNA technology [See, for example, Genetics Institute, PCT application W085/01961; Genentech European Patent application 160,457; Chiron European Patent application 150,735; J. J. Toole et al., "Molecular Cloning Of a cDNA Encoding Human Antihaemophilic Factor" Nature, 312, pp. 342-47 (1984); and W. I. Wood et al., Nature, 312, pp. 330-37 (1984)]. However, such attempts have proven to be less successful than had been hoped. This is partially due to the fact that the recombinantly produced 2332 amino acid factor VIII:C chain is subject to proteolytic cleavage at many positions. It is also due to difficulties in producing recombinant factor VIII:C in sufficiently high yields.

SUMMARY OF THE INVENTION

The present invention solves the problems referred to above by providing DNA sequences which encode modified factor VIII:C and modified factor VIII:C-like polypeptides. These DNA sequences code for polypeptides which are produced in approximately twenty-times higher yields than previous recombinantly produced factor VIII:C and are more easily purified into biochemically pure mature factor VIII:C.

According the present invention, DNA sequences coding for modified factor VIII:C are produced and expressed in high yields. As will be apparent from the disclosure and examples to follow, the modified factor VIII:C and modified factor VIII:C-like polypeptides of this invention are characterized by deletions removing a major part of the maturation polypeptide of factor VIII:C. The DNA sequences in

our preferred embodiment have a deletion of substantially all of the nucleotides coding for the maturation polypeptide. Our most preferred embodiment contains a deletion of all the DNA sequence coding for the maturation polypeptide. On expression of our DNA sequences, the heavy chain of mature factor VIII:C is linked directly to the light chain. Following a one-nick proteolytic event, the mature form of factor VIII:C is generated.

Finally, the present invention provides various anti-haemophilic compositions containing modified factor VIII:C and modified factor VIII:C-like polypeptides produced by the DNA sequences of this invention, and various methods of using those compositions in haemophilia treatment-therapy of acute or prolonged bleeding in haemophilia A.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts a restriction map of the factor VIII:C cDNA.

Figure 2 is a schematic depiction of the construction of the recombinant DNA molecule with the QD deletion.

Figures 3A and 3B depict a schematic representation of the construction of the recombinant DNA molecule with the RE deletion.

Figure 4 depicts a restriction endonuclease map of the RE deletion inserted into the mammalian cell expression vector pBG312 indicating the positions of the SV40 origin of replication/enhancer, the adenovirus major late promoter, the factor VIII:C cDNA with the RE deletion, the 3' untranslated region of the factor VIII:C mRNA, and the polyadenylation site.

Figure 5 depicts the results of an S1 analysis of Factor VIII:C mRNA isolated from transfected BMT10 cells.

Figure 6 depicts the results of a Southern analysis of plasmid DNA isolated from transfected BMT10 cells.

Figure 7 depicts the published DNA and amino acid sequence of factor VIII:C (EPO application 160,457).

DETAILED DESCRIPTION OF THE INVENTION

In order that the invention herein described may be more fully understood, the following detailed description is set forth.

In the description the following terms are employed:

Nucleotide--A monomeric unit of DNA or RNA consisting of a sugar moiety (pentose), a phosphate, and a nitrogenous heterocyclic base. The base is linked to the sugar moiety via the glycosidic carbon (1' carbon of the pentose) and that combination of base and sugar is called a nucleoside. The base characterizes the nucleotide. The four DNA bases are adenine ("A"), guanine ("G"), cytosine ("C"), and thymine ("T"). The four RNA bases are A, G, C, and uracil ("U").

DNA Sequence--A linear array of nucleotides connected one to the other by phosphodiester bonds between the 3' and 5' carbons of adjacent pentoses.

Codon--A DNA sequence of three nucleotides (a triplet) which encodes through mRNA an amino acid, a translation start signal or a translation termination signal. For example, the nucleotide triplets TTA, TTG, CTT, CTC, CTA and CTG encode for the amino acid leucine ("Leu"), TAG, TAA and TGA are translation stop signals and ATG is a translation start signal.

Amino Acid--A monomeric unit of a peptide, polypeptide or protein. The twenty amino acids are: phenylalanine ("Phe" or "F"), leucine ("Leu", "L"),

isoleucine ("Ile", "I"), methionine ("Met", "M"), valine ("Val", "V"), serine ("Ser", "S"), proline ("Pro", "P"), threonine ("Thr", "T"), alanine ("Ala", "A"), tyrosine ("Tyr", "Y"), histidine ("His", "H"), glutamine ("Gln", "Q"), asparagine ("Asn:N"), lysine ("Lys:K"), aspartic acid ("Asp", "D"), glutamic acid ("Glu", "E"), cysteine ("Cys", "C"), tryptophane ("Trp", "W"), arginine ("Arg", "R") and glycine ("Gly", "G").

Reading Frame--The grouping of codons during the translation of mRNA into amino acid sequences. During translation the proper reading frame must be maintained. For example, the DNA sequence GCTGGTTGTAAG may be expressed in three reading frames or phases, each of which affords a different amino acid sequence:

GCT GGT TGT AAG--Ala-Gly-Cys-Lys G
CTG GTT GTA AG--Leu-Val-Val GC TGG
TTG TAA G--Trp-Leu-(STOP)

Polypeptide--A linear array of amino acids connected one to the other by peptide bonds between the α-amino and carboxy groups of adjacent amino acids.

Genome--The entire DNA of a cell or a virus. It includes inter alia the structural gene coding for the polypeptides of the substance, as well as operator, promoter and ribosome binding and interaction sequences, including sequences such as the Shine-Dalgarno sequences.

Gene--A DNA sequence which encodes through its template or messenger RNA ("mRNA") a sequence of amino acids characteristic of a specific polypeptide.

Transcription--The process of producing mRNA from a gene or DNA sequence.

Translation--The process of producing a polypeptide from mRNA.

Expression--The process undergone by a gene or DNA sequence to produce a polypeptide. It is a combination of transcription and translation.

Plasmid--A nonchromosomal double-stranded DNA sequence comprising an intact "replicon" such that the plasmid is replicated in a host cell. When the plasmid is placed within a unicellular organism, the characteristics of that organism may be changed or transformed as a result of the DNA of the plasmid. For example, a plasmid carrying the gene for tetracycline resistance (TET^R) transforms a cell previously sensitive to tetracycline into one which is resistant to it. A cell transformed by a plasmid is called a "transformant".

Phage or Bacteriophage--Bacterial virus, many of which consist of DNA sequences encapsidated in a protein envelope or coat ("capsid").

Cloning Vehicle--A plasmid, phage DNA, cosmid or other DNA sequence which is able to replicate in a host cell, characterized by one or a small number of endonuclease recognition sites at which such DNA sequences may be cut in a determinable fashion without attendant loss of an essential biological function of the DNA, e.g., replication, production of coat proteins or loss of promoter or binding sites, and which contains a marker suitable for use in the identification of transformed cells, e.g., tetracycline resistance or ampicillin resistance. A cloning vehicle is often called a vector.

Cloning--The process of obtaining a population of organisms or DNA sequences derived from one such organism or sequence by asexual reproduction.

Recombinant DNA Molecule or Hybrid DNA--A molecule consisting of segments of DNA from different genomes which have been joined end-to-end outside of living cells and able to be maintained in living cells.

Expression Control Sequence--A sequence of nucleotides that controls and regulates expression of genes when operatively linked to those genes. They include the lac system, the β-lactamase system, the trp system, the tac and trc systems, the major operator and promoter regions of phage λ, the control region of fd coat protein, the early and late promoters of SV40, promoters derived from polyoma virus and adenovirus, metallothionine promoters, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase, e.g., Pho5, the promoters of the yeast α-mating factors, and other sequences known to control the expression of genes of prokaryotic or eukaryotic microbial cells and their viruses or combinations thereof.

Factor VIII:C --A polypeptide having the amino acid sequence of Figure 7, and upon maturation and activation, being capable of functioning as co-factor for the factor IXa-dependent maturation of factor X in the blood coagulation cascade. As used in this application, factor VIII:C includes the glycoproteins also known as factor VIII procoagulant activity protein, factor VIII-clotting activity, antihemophilic globulin (AHG), antihemophilic factor (AHF), and antihemophilic factor A [see W. J. Williams et al., Hematology, pp. 1056, 1074 and 1081].

Maturation Polypeptide --The maturation polypeptide of factor VIII:C is made up of the 908 amino acids from amino acid Ser (741) to amino acid Arg (1648) (see Figure 7). Maturation of factor VIII:C is initiated with a cleavage between amino acids 1648 and 1649 (which produces a C-terminal light chain) followed by a series of nicks which produce the mature N-terminal heavy chain.

Mature Factor VIII:C --As used in this application, mature factor VIII:C is composed of an N-terminal heavy chain (Ala 1- Arg 740) linked to a C-terminal light chain (Glu 1649-Tyr 2332) through an alkaline metal bridge, such as calcium (Figure 7).

Modified Factor VIII:C -- As used in this application, "modified factor VIII:C" refers to polypeptides characterized by a deletion of a major portion of the maturation polypeptide of factor VIII:C. For example, where the entire maturation polypeptide has been deleted, "modified factor VIII:C" includes proteins that comprise the N-terminal mature heavy chain and the C-terminal mature light chain of factor VIII:C linked together as a single chain.

Modified Factor VIII:C-Like Polypeptide -- As used in this application, "modified factor VIII:C-like polypeptide" includes proteins having the biological activity of modified factor VIII:C. It also includes proteins having an amino terminal methionine, e.g., f-Met-factor VIII:C, and proteins that are characterized by other amino acid deletions, additions or substitutions so long as those proteins substantially retain the biological activity of modified factor VIII:C.

"Modified factor VIII:C-like polypeptides" within the above-definition also includes natural allelic variations that may exist and occur from individual to individual. Furthermore, it includes modified factor VIII:C-like polypeptides whose degree and location of glycosylation, or other post-translation modifications, may vary depending on the cellular environment of the producing host or tissue.

The present invention relates to processes for the production of modified factor VIII:C and modified factor VIII:C-like polypeptides. More particularly, it provides DNA sequences which permit the production of modified factor VIII:C and modified factor VIII:C-like polypeptides in high yields, in appropriate hosts. Polypeptides produced by the DNA sequences of this invention are useful in the clinical treatment of haemophilia A.

As compared to factor VIII:C, the modified factor VIII:C produced by the DNA sequences of this invention lack a major portion of the maturation polypeptide of factor VIII:C. The DNA sequences of the present invention surprisingly express modified factor VIII:C in much higher yields than DNA sequences coding for factor VIII:C itself.

While not wishing to be bound by theory, we believe that the DNA sequences of the present invention produce modified factor VIII:C in high yields because of the absence of most or all of the maturation polypeptide. For example, the mRNA for the modified gene may be translated more efficiently, because the RNA coding for the long maturation polypeptide does not have to be translated. In addition, while factor VIII:C has many proteolytic targets which may be attacked while the polypeptide is in the cell, the modified factor VIII:C is less subject to such proteolytic attack because it lacks the proteolytic targets within the maturation polypeptide. Furthermore, when the maturation polypeptide is absent, 19 of the 25 N-linked glycosylation sites of native factor VIII:C are deleted, leaving only six N-liked glycosylation

sites on the modified polypeptide (three on the heavy chain and three on the light chain). Apparently, because there are fewer sites to be glycosylated, production and purification of the modified factor VIII:C is simplified.

In the processes of this invention, we modify the DNA sequence encoding factor VIII:C to delete from it a major portion of the DNA sequence encoding the maturation polypeptide. Having prepared a DNA sequence carrying the desired deletion we employ it in a variety of expression vectors and hosts to produce modified factor VIII:C encoded by it. For example, any of a wide variety of expression vectors are useful in expressing the modified factor VIII:C coding sequences of this invention. It also should be understood that DNA sequences encoding a modified factor VIII:C-like polypeptide can be similarly produced in accordance with this invention.

Useful expression vectors include, for example, vectors consisting of segments of chromosomal, non-chromosomal and synthetic DNA sequences, such as various known derivatives of SV40, known bacterial plasmids, e.g., plasmids from E.coli including col E1, pCR1, pBR322, pMB9 and their derivatives, wider host range plasmids, e.g., RP4, phage DNAs, e.g., the numerous derivatives of phage λ, e.g., NM 989, and other DNA phages, e.g., M13 and Filamenteous single stranded DNA phages, yeast plasmids such as the 2μ plasmid or derivatives thereof, and vectors derived from combinations of plasmids and phage DNAs, such as plasmids which have been modified to employ phage DNA or other expression control sequences. In the preferred embodiments of this invention, we employ pBG312, a pBR327-related vector [R. Cate et al., Cell, 45, pp. 685-98 (1986)].

In addition, any of a wide variety of expression control sequences -- sequences that con-

trol the expression of a DNA sequence when operatively linked to it -- may be used in these vectors to express the DNA sequence of this invention. Such useful expression control sequences, include, for example, the early and late promoters of SV40, the lac system, the trp system, the TAC or TRC system, the major operator and promoter regions of phage λ, the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase, e.g., Pho5, the promoters of the yeast α-mating factors, and other sequences known to control the expression of genes of prokaryotic or eukaryotic cells or their viruses, and various combinations thereof. In the preferred embodiment of this invention, we employ adenovirus-2 major late promoter expression control sequences.

A wide variety of host cells are also useful in producing the modified factor VIII:C of this invention. These hosts may include well known eukaryotic and prokaryotic hosts, such as strains of E.coli, Pseudomonas, Bacillus, Streptomyces, fungi such as yeasts, and animal cells, such as CHO cells, African green monkey cells, such as COS1, COS7, BSC1, BSC40, and BMT10, and human cells and plant cells in tissue culture. In the preferred embodiments of this invention, we prefer BMT10 African green monkey cells.

It should of course be understood that not all vectors and expression control sequences will function equally well to express the modified DNA sequences of this invention and to produce our modified factor VIII:C. Neither will all hosts function equally well with the same expression system. However, one of skill in the art may make a selection among these vectors, expression control sequences and hosts without undue experimentation without departing from the scope of this invention. For

example, in selecting a vector, the host must be considered because the vector must replicate in it. The vector's copy number, the ability to control that copy number, and the expression of any other proteins encoded by the vector, such as antibiotic markers, should also be considered.

In selecting an expression control sequence, a variety of factors should also be considered. These include, for example, the relative strength of the system, its controllability, and its compatibility with the DNA sequence encoding the modified factor VIII:C of this invention, particularly as regards potential secondary structures. Hosts should be selected by consideration of their compatibility with the chosen vector, the toxicity of our modified factor VIII:C to them, their secretion characteristics, their ability to fold proteins correctly, their fermentation requirements, and the ease of the purification of our modified factor VIII:C from them and safety.

Within these parameters one of skill in the art may select various vector/expression control system/host combinations that will produce useful amounts of our modified factor VIII:C on fermentation. For example, in one preferred embodiment of this invention, we use an pBG312 vector, with an adenovirus 2 major late promoter expression system in BMT10 African green monkey cells.

The modified factor VIII:C and modified factor VIII-like polypeptides produced according to this invention may be purified by a variety of conventional steps and strategies. Useful purification steps include those used to purify natural and recombinant factor VIII:C [see, for example, L.-O. Andersson et al., PNAS (USA), 83, pp. 2979-83 (1986)].

After purification the modified factor VIII:C and modified factor VIII:C-like polypeptides

of this invention are useful in composition and methods for treatment of haemophilia A and in a variety of agents useful in treating uncontrolled bleeding.

While the modified factor VIII:C and modified factor VIII:C-like polypeptides of this invention may be administered in such compositions and methods in the form in which they are produced, as single chain polypeptides, it should also be understood that it is within the scope of this invention to administer the modified factor VIII:C after subjecting it to proteolytic cleavage. For example, modified factor VIII:C can be cleaved in vitro, into the heavy chain and light chain of mature factor VIII:C and linked with a calcuim or other alkaline metal bridge, before, during or after purification.

The modified factor VIII:C and modified factor VIII:C-like polypeptides of this invention may be formulated using known methods to prepare pharmaceutically useful compositions. Such compositions also will preferably include conventional pharmaceutically acceptable carriers and may include other medicinal agents, carriers, adjuvants, excipients, etc., e.g., human serum albumin or plasma preparations. See, e.g., Remington's Pharmaceutical Sciences (E. W. Martin). The resulting formulations will contain an amount of modified factor VIII:C effective in the recipient to treat uncontrolled bleeding. Administration of these polypeptides, or pharmaceutically acceptable derivatives thereof, may be via any of the conventional accepted modes of administration of factor VIII. These include parenteral, subcutaneous, or intravenous administration.

The compositions of this invention used in the therapy of haemophilia may also be in a variety of forms. The preferred form depends on the intended mode of administration and therapeutic application.

The dosage and dose rate will depend on a variety of factors for example, whether the treatment is given to an acutely bleeding patient or as a prophylactic treatment. However, the factor VIII:C level should be high enough to prevent hemorrhage and promote epithelialization [see discussion in Williams, Hematology, pp. 1335-43].

In order that this invention may be better understood, the following example is set forth. This example is for purposes of illustration only and is not to be construed as limiting the scope of the invention.

EXAMPLE

We have constructed cDNA sequences which encode modified factor VIII:C molecules having a deletion of a major part of or all of the maturation polypeptide. To test the limits of our invention, we also constructed a cDNA sequence which encodes a polypeptide having a deletion of more than just the maturation polypeptide of factor VIII:C.

A. ASSEMBLY OF THE FULL-LENGTH FACTOR VIII:C cDNA

Referring now to Figure 1, we have presented therein a restriction enzyme map of the factor VIII:C cDNA based upon the published sequence [W. I. Wood et al., Nature, 312, pp. 330-37 (1984); (Figure 7)]. The bar represents the coding sequence. Below the restriction enzyme map we have depicited the amino-terminal heavy chain of mature factor VIII:C attached by a calcium bridge to the carboxy-terminal light chain of mature factor VIII:C. Below the protein model on a bar congruent to the restriction enzyme map we have indicated the oligonucleotide probes (indicated with asterisks) which we used to screen human placenta, liver, and kidney cDNA libraries.

These libraries were made using oligo (dT) as first-strand primer and λgt10 as vector.

On this second bar are also located the oligonucleotide primers (left-arrows) which we used to initiate first-strand cDNA synthesis using human kidney mRNA as template. We made these single-stranded cDNA sequences double-stranded by the technique of Gubler and Hoffman [U. Gubler, and B. J. Hoffman, Gene, 25, pp. 263-69 (1983)]. We cloned them at the dC-tailed EcoRV site in pBR322. We then screened this plasmid-based kidney cDNA library with oligonucleotide probes located on the bar 5' to the oligonucleotide primers.

Below the primer/probe bar in Figure 1, we have displayed a collection of partial-length factor VIII:C cDNA and genomic subclones, which we isolated from these libraries. Together these encode the full-length cDNA gene. More information about these clones is presented below, in Table 1.

The dosage and dose rate will depend on a variety of factors for example, whether the treatment is given to an acutely bleeding patient or as a prophylactic treatment. However, the factor VIII:C level should be high enough to prevent hemorrhage and promote epithelialization [see discussion in Williams, Hematology, pp. 1335-43].

In order that this invention may be better understood, the following example is set forth. This example is for purposes of illustration only and is not to be construed as limiting the scope of the invention.

EXAMPLE

We have constructed cDNA sequences which encode modified factor VIII:C molecules having a deletion of a major part of or all of the maturation polypeptide. To test the limits of our invention, we also constructed a cDNA sequence which encodes a polypeptide having a deletion of more than just the maturation polypeptide of factor VIII:C.

A. ASSEMBLY OF THE FULL-LENGTH FACTOR VIII:C cDNA

Referring now to Figure 1, we have presented therein a restriction enzyme map of the factor VIII:C cDNA based upon the published sequence [W. I. Wood et al., Nature, 312, pp. 330-37 (1984); (Figure 7)]. The bar represents the coding sequence. Below the restriction enzyme map we have depicited the amino-terminal heavy chain of mature factor VIII:C attached by a calcium bridge to the carboxy-terminal light chain of mature factor VIII:C. Below the protein model on a bar congruent to the restriction enzyme map we have indicated the oligonucleotide probes (indicated with asterisks) which we used to screen human placenta, liver, and kidney cDNA libraries.

These libraries were made using oligo (dT) as first-strand primer and λgt10 as vector.

On this second bar are also located the oligonucleotide primers (left-arrows) which we used to initiate first-strand cDNA synthesis using human kidney mRNA as template. We made these single-stranded cDNA sequences double-stranded by the technique of Gubler and Hoffman [U. Gubler, and B. J. Hoffman, Gene, 25, pp. 263-69 (1983)]. We cloned them at the dC-tailed EcoRV site in pBR322. We then screened this plasmid-based kidney cDNA library with oligonucleotide probes located on the bar 5' to the oligonucleotide primers.

Below the primer/probe bar in Figure 1, we have displayed a collection of partial-length factor VIII:C cDNA and genomic subclones, which we isolated from these libraries. Together these encode the full-length cDNA gene. More information about these clones is presented below, in Table 1.

TABLE 1

COMPENDIUM OF FACTOR VIII:C
GENOMIC AND cDNA CLONES

isolated from a genomic library constructed with cosmid pTCF [Grosveld, F.G. et. al., Nucleic Acids Research, 10, pp. 6715-6732 (1982)]

subclone	length	tissue
pUC19.2874	2874 bp	48,XXXX human lymphoblast

isolated from oligo(dT)-primed λgt10 cDNA libraries

clone	length	probe hybridization
1.7977 (placenta)	1728	79+, 77+
2.73 (liver)	~700	73+
4.73 (kidney)	~220	73+

isolated from a 85, 86-primed Gubler-Hoffman kidney cDNA library

clone	length	probe hybridization
1.82	~1200	82+, 79-, 77-
2.82	~1200	82+, 79-, 77-
3.7573	~2700	74-, 75+, 73+
4.7573	~2700	74-, 75+, 73+
6.7573	~2700	74-, 75+, 73+
10.797783	>1263	82-, 79+, 77+, 83+
11.797783	>1263	82-, 79+, 77+, 83+
12.797783	>1263	82-, 79+, 77+, 83+
13.797783	>1263	82-, 79+, 77+, 83+

isolated from a 75, 77-primed Gubler-Hoffman kidney cDNA library

clone	length	probe hybridization
7.7475	~2700	74+, 75+

Prior to assembling the full-length cDNA gene, we constructed two intermediate plasmids. This was necessary because of the excessive length of the factor VIII:C cDNA. For our first preliminary construction we isolated a fragment from clone 7.7475 extending from the PstI site at 5163 to the PstI site at 5755. We inserted this fragment into clone 4.7573 at the PstI site at 5755 thereby extending clone 4.7573. This PstI site is shared by the inserts of both clone 7.7475 and clone 4.7573. By extending clone 4.7573 in this manner, we provided a unique NdeI site at 5522 in the insert of this derivative of clone 4.7573. We needed to create this Nde site because we needed a unique site at which to extend the length of this insert at its 5' end.

As a second preliminary construction we introduced a polynucleotide linker in clone 2.82 at a location immediately 5' to the translation start codon of the signal sequence of factor VIII:C. The insert of clone 2.82 is at the EcoRV site of pBR322 and its orientation is opposite to that of tetracycline resistance. The 5' endpoint of the insert in clone 2.82 is at -133 in the 5' untranslated leader sequence. We cleaved clone 2.82 at the SalI site in tetracycline resistance and at the SacI site in the sequence encoding the signal peptide in the insert of clone 2.82 and inserted the synthetic duplex

```
          H                                   BH
          SAI     N     N                     BSGNSS
          ACN     R     C                     APISAS
          LCC     U     O                     N1APCT
          112     1     1                     221211
                                              /////
          GTCGACTCGCGACCATGGATGCAAATAGAGCTC
        1 ---------+-------+---------+---- 33
          CAGCTGAGCGCTGGTACCTACGTTTATCTCGAG
                             MetGlnIleGluLeu
```

This ligation resulted in the introduction of a SalI-NruI-NcoI polylinker immediately 5' to the start

codon which initiates translation of the signal sequence of factor VIII:C. These three restriction enzymes do not cleave the full-length factor VIII:C cDNA gene.

With these two intermediate constructions available, we assembled the full-length factor VIII:C cDNA in a six-fragment ligation reaction (bottom Figure 1). It was necessary to create the full length DNA in this manner because we never isolated the full DNA in one single clone. We isolated fragment 1 from the above-described derivative of clone 2.82. Fragment 1 extended from SalI in the polylinker to AvaI at 731. We isolated Fragment 2 from the insert in the λgt10 recombinant 1.7977. Fragment 2 extended from AvaI at 731 to EcoRI at 2289. Fragment 3 derived from the subclone pUC19.2874 of a genomic cosmid recombinant; it extended from EcoRI at 2289 to BamHI at 4743. Fragment 4 was isolated from clone 7.74575, starting from the BamHI site at 4743 and extending to the NdeI site at 5522. We isolated fragment 5 from the above-described derivative of clone 4.7573. Fragment 5 extended from NdeI at 5522 to NcoI at 7991. Fragment 6 is an assembly vector containing an E.coli replication origin and selectable marker for ampicillin resistance. We isolated Fragment 6 from pAT.SV2.tPA, a gift from Richard Fisher. This is a plasmid in which the transcription of the tPA gene is under the control of the SV40 early promoter. We digested pAT.SV2.tPA with SalI which cleaves within the tetracycline resistance marker, and with NcoI which cleaves within the SV40 early region.

Of the 96 recombinants we analyzed, 32 contained all five factor VIII:C restriction fragments. We determined the DNA sequence of one of these clones, and we identified two changes with respect to the published sequence. One is a CTG to

CTA change at Leu 242 and the other is a TTC to CTC change at amino acid residue 1880 (Phe to Leu) (compare Figure 7).

B. INSERTION OF THE FULL-LENGTH cDNA INTO A MAMMALIAN CELL EXPRESSION VECTOR

We excised the full-length factor VIII:C cDNA gene from the assembly vector by digestion with NcoI. We then treated the resultant NcoI restriction fragment with nuclease S1 to create a blunt end. We ligated this fragment to SmaI-digested pBG312. pBG312 is an animal cell expresion vector whose construction has been described elsewhere [R. Cate et al., Cell, 45, pp. 685-98 (1986)]. The sequence of BG312, from EcoRI to BamHI has (clockwise): a SV40 replication origin; an adenovirus-2 major late promoter and complete tripartite leader [S. Zain et al., Cell, 16, pp. 851-61 (1979)]; a hybrid splice signal consisting of an adenovirus-5 splice donor and an immunoglobulin variable region gene splice acceptor [R. J. Kaufman, and P. A. Sharp, J. Mol. Biol., 159, pp. 601-21 (1982)]; a polylinker containing sites for HindIII, XhoI, EcoRI, SmaI, NdeI, SstI, and BglII; the SV40 small t antigen intron flanked by its splice donor and acceptor; and the SV40 early polyadenylation site.

We verified the DNA sequence across the junction between the polylinker of pBG312 and the cDNA gene encoding factor VIII:C including the signal sequence for two independent clones: 8.1 and 8.2. Clone 8.1 differs from 8.2 in the 3' untranslated region; T 7806 is fused to the SmaI site of pBG312 in clone 8.1 instead of the C of the NcoI site at 7990 in clone 8.2. In addition, we isolated another clone, in which the fusion of the cDNA gene encoding factor VIII:C to pBG312 had occured within the sequence encoding the signal peptide of factor

VIII:C. This clone, which we named signal-minus, provided a negative control for our transient expression assays, described below.

C. CONSTRUCTION OF GLN 744 - ASP 1563 (ABBREVIATED QD) DELETION

In this section we demonstrate how we created the QD deletion which removes a portion of DNA sequence coding on expression for the maturation polypeptide (amino acids 741-1648). The QD deletion retains approximately 90 amino acids of the maturation polypeptide (four amino acides at the N-terminal end of the maturation polypeptide and 86 amino acids at its carboxy terminal end).

Referring now to Figure 2, we depict therein the construction of the QD deletion. We partially digested one aliquot of the expression plasmid for the full-length factor VIII:C gene with EcoRI. This endonuclease cleaves between the codons for Gln 744 and Asn 745. We removed the 5'AATT overhang with nuclease S1, and then subjected the plasmid to complete digestion with PvuI within the ampicillin resistance gene. We partially digested another aliquot with BamH1, which cleaves between the codons for Leu 1562 and Asp 1563 (see Figure 7). We filled out the 5'GATC overhang with the Klenow fragment, and again digested the plasmid with PvuI within amp. We then combined the two mixtures of fragments and ligated them with T4 DNA ligase. A BamH1 site between the codons for Gln 744 and Asp 1563 was created in this fusion.

The modified polypeptide produced on expression as a result of the QD deletion lacks 818 amino acids from within the 908 amino-acid maturation polypeptide, leaving 4 amino acids C-terminal to the carboxy terminus of the mature heavy chain, Arg 740, and leaving 86 amino acids N-terminal to the amino

terminus of the light chain, Glu 1649 (Figure 7). The 908 amino-acid maturation polypeptide is thus replaced by a 90 amino-acid maturation polypeptide, with the protease substates for both initial maturation of the primary translation polypeptide and subsequent maturation of the heavy chain remaining intact.

D. CONSTRUCTION OF THE ARG 740 - GLU 1649 (ABBREVIATED RE) DELETION

We demonstrate in this section how we created the RE deletion, which removes the entire DNA sequence coding for the maturation polypeptide.

Referring now to Figures 3A and B, we show how we obtained this RE deletion fusion in two steps. In the first step we ligated four fragments which resulted in an intermediate plasmid. These four fragments were:

(1) the 462 bp fragment, obtained by digesting the expression plasmid for the full-length gene with HindIII between the codons for Arg 740 and Ser 741, removing the 5' AGCT with nuclease S1, and subsequently digesting with KpnI which cleaves uniquely between the codons for Tyr 586 and Leu 587.

(2) the synthetic oligonucleotide duplex fragment

```
5'pGAA ATA ACT CGT ACT ACT CTT CAG TCA
   CTT TAT TGA GCA TGA TGA GAA GTC AGT CTA Gp 5'
   Glu Ile Thr Arg Thr Thr Leu Gln Ser Asp
   1649                                1657
```

(3) the 135 bp fragment obtained by digesting the expression plasmid for the full-length gene first with Sau3A; we isolated the 411 bp fragment which resulted from Sau3A digestion between the codons for Ser 1657 and Asp 1658 and between the codons for Glu 1794 and Asp 1795. Then, we digested the 411 bp fragment with PstI which cleaves between the codons

for Ala 1702 and Val 1703, to obtain the 135 bp 5' fragment.

(4) pUC18 digested with KpnI and PstI.

We then isolated a fragment encoding the RE fusion from this intermediate plasmid. To do this, we digested the intermediate plasmid generated in the four-fragment ligation with Asp718 and PstI. The fragment encoding the RE fusion was used to replace the corresponding fragment in the expression plasmid for the QD fusion. We ligated the resultant 624 bp fragment encoding the RE fusion to the mixture of fragments which we obtained by first completely digesting the expression plasmid for the QD internal deletion at the unique Asp718 site, next dephosphorylating the 5' GTAC overhang with calf intestinal phosphatase, and then partially digesting the plasmid with PstI.

Referring now to Figure 4, we depict therein a map of the RE deletion inserted into pBG312. In the modified polypeptide produced on expression the 908 amino-acid maturation polypeptide is entirely removed. The novel polypeptide produced by this recombinant molecule cell is secreted, and may be purified as a single chain, i.e., the heavy chain is linked directly to the light chain. Because the Arg 1648 - Glu 1649 peptide bond which is normally cleaved during the initial nicking of the full-length primary translation product is preserved in this deletion, the primary translation product for this internal deletion is nicked by the same protease that initiates nicking of the full-length primary translation product, thus producing directly the mature form of the heavy chain of factor VIII:C. Our Western blot analysis (data not shown) confirms that the RE modified factor VIII:C encodes a single chain molecule which is then processed into a 90K heavy chain and an 80K light

chain in the culture medium. The resultant light chain possesses the peptide from Glu 1649 to Arg 1689 that binds the two-chain complex to von Willebrand protein. For this reason, this recombinant product, when secreted from a mammalian cell, will bind to the von Willebrand protein present in cell culture fluid. Similarly, when injected, it will complex to and circulate with plasma von Willebrand protein. Upon thrombin cleavage at Arg 1689 - Ser 1690, the two-chain mature factor VIII:C will be activated and will dissociate from von Willebrand protein and assemble into its ternary complex with factor IXa and factor X on a platelet surface.

E. CONSTRUCTION OF THE ARG 740 - SER 1690 (ABBREVIATED RS) DELETION

In order to test the outer limits of these deletions, we constructed a plasmid which codes for a polypeptide with a deletion of more than the maturation polypeptide alone (i.e., we deleted the DNA sequence which codes on expression for the forty-one amino acids at the N-terminal end of the light chain of mature factor VIII:C).

We constructed this RS fusion with the two-step strategy described above for the RE fusion. Our first step was a three-fragment ligation resulting in an intermediate plasmid. The three fragments which we ligated were:

(1) the 462 bp fragment, obtained by digesting the expression plasmid for the full-length gene with HindIII between the codons for Arg 740 and Ser 741, removing the 5' AGCT with nuclease S1, and subsequently digesting with KpnI which cleaves uniquely between the codons for Tyr 586 and Leu 587.

(2) the synthetic oligonucleotide duplex fragment:

5' pAGC TTT CAA AAG AAA ACA CGA CAC TAT TTT ATT GCT GCA
TCG AAA GTT TTC TTT TGT GCT GTG ATA AAA TAA CGp 5'
Ser Phe Gln Lys Lys Thr Arg His Tyr Phe Ile Ala Ala
1690 1702

(3) pUC18 digested with KpnI and PstI.

In this fusion, we recreated the HindIII site between the codons for Arg 740 and Ser 741 (now Ser 1690).

We isolated a fragment encoding the RS fusion from this intermediate plasmid and used this fragment in our second step to replace the corresponding fragment in the expression plasmid for the QD fusion. In this second step, we isolated a 501 bp fragment encoding the RS fusion. We digested the intermediate plasmid with Asp718 and PstI and isolated the fragment encoding the RS fusion. We then used the strategy described above for the RE fusion to replace the related fragment in the expression plasmid for the QD fusion with the 501 bp fragment.

In addition to removing the entire maturation polypeptide, the RS deletion removes DNA coding for the Glu 1649 - Arg 1689 peptide, the putative von Willebrand binding domain. For this reason this recombinant molecule will not attach to circulating von Willebrand protein when it is secreted from an animal cell into culture fluid or when it is injected into a recipient.

F. TRANSFECTION OF AFRICAN GREEN MONKEY KIDNEY CELLS

We transfected BMT10 cells [R. D. Gerard and Y. Gluzman, Mol. Cell. Biol., 5, pp. 3231-40 (1985)] with the supercoiled expression plasmid. We used the DEAE-dextran technique [L. M. Sompayvac and K. J. Danna, PNAS, 78, pp. 7575-78 (1981)] and chloroquine [H. Luthman and G. Magnusson, Nucleic Acids Research, 11, pp. 1295-1308 (1983)] to trans-

fect the cells. Transfectants are known to replicate the input expression plasmid to high copy number because SV40 T antigen is inducibly supplied in trans by BMT10 cells and binds to the SV40 origin of replication linked to the modified factor VIII:C gene in the expression plasmids. However, this technique is inefficient because, typically, only several percent of the transfected cells will actually incorporate DNA.

The transfectants will secrete modified factor VIII:C for up to 120 hours. For most experiments, the cm^2/ml ratio is approximately 5.5; that is, a confluent monolayer of BMT10 transfectants in a 100 mm Petri dish (55 cm^2) is covered with 10 ml culture fluid.

G. FACTOR VIII:C ACTIVITY ASSAY

We assayed the signal-minus, 8.1, QD, RE and RS expression constructs for factor VIII:C production after transfection in duplicate into BMT10 cells. We used a 96-well plate adaptation of KabiVitrum's Coatest® Factor VIII:C. One petri dish was used to prepare RNA for S1 analysis and the other petri dish was used to prepare Hirt DNA used in our Southern analysis. After 120 hours of incubation we assayed the cell culture fluids for factor VIII:C activity. We expressed our results in terms of % plasma level, where plasma factor VIII:C concentration is approximately 200 ng/ml.

In repeated transfections, both the signal-minus construct (negative control) and the RS deletion have shown no detectable factor VIII:C activity. This may be explained by the deletion of the von Willebrand protein binding domain in the RS deletion.

In the 120 hour experiment analyzed below, cells transfected with the full-length gene produced approximately 5% of the activity observed with both the QD and the RE deletions. The activity observed

with the QD deletion was 1.46% plasma level and that for the RE deletion was 1.30% plasma level.

Thus, we observed that BMT10 cells transfected with the QD and RE deletions produce at least 20 times more factor VIII:C than cells transfected with the full-length gene.

H. NUCLEASE S1 ANALYSIS OF FACTOR VIII:C mRNA

In order to determine the levels of mRNA in each construction, we conducted a nuclease S1 analysis. This assay assists in the determination of the reason for the increased level of expression in our QD and RE deletions.

We isolated RNA from 100 mm Petri dish cultures of BMT10 cells 120 hours after transfection, using the unpublished method of W. Schleuning and J. Bertonis. Briefly, according to this method, we lysed BMT10 cells with 3 ml of 50 mM Tris-HCl (pH 7.5) - 5 mM EDTA - 1% SDS containing 100 µg/ml proteinase K for 20 minutes at 37°C. We transferred the lysate to a 50 ml conical tube containing 3 ml of phenol and then mechanically sheared the DNA for 15 seconds at high speed in a Polytron (Brinkmann Instruments). We extracted the aqueous phase with ether and adjusted it to 0.25 NaCl. We precipitated the nucleic acid fraction at 4°C, by the addition of an equal volume of isopropanol, collected it by centrifugation and redissolved it in 3 ml of water. We selectively precipitated RNA overnight at 4°C, by adjusting the solution to 2.8 M LiCl.

We determined the amount of modified factor VIII:C mRNA for each construction. We isolated probes for the S1 analysis by digesting the QD expression plasmid with EspI. We labelled the 5' ends of the EspI fragments with [γ-^{32}P]ATP and T4 polynucleotide kinase, and annealed 10 µg RNA to 5000 cpm of the ^{32}P-antisense strand of the 477 nucleotide EspI frag-

ment isolated on a 5% strand separation gel [A. M. Maxam and W. Gilbert, Methods In Enzymology, 65, pp. 499-560 (1980)]. We incubated the RNA overnight at 48°C in 10 µl 80% deionized formamide - 400 mM NaCl - 40 mM PIPES (pH 6.4) - 1 mM EDTA. The hybrid molecules were then digested for 60 minutes at 37°C by adding 190 µl nuclease S1 at a concentration of 100 units/ml in 0.28 M NaCl - 50 mM NaOAc (pH 4.6) - 4.5 mM ZnSO4. We terminated the digestion by adding EDTA to 10 mM and extracting with phenol. We denatured the protected fragments and subjected them to electrophoresis on a 5% strand separation gel. We exposed the dried gel to Kodak XAR-5 X-ray film backed by a Lightning-Plus intensifying screen (Dupont) overnight at -70°C. The 477 nucleotide EspI fragment has one end within the hybrid intron spliced out from the 5' untranslated region of the factor VIII:C mRNA [R. J. Kaufman and P. A. Sharp, J. Mol. Biol., 159, pp. 601-21 (1981)] and the other end within the codon for Ala 62 (Figure 4).

We detected modified factor VIII:C mRNA by protecting a single-stranded 300 nucleotide DNA fragment from digestion. The experiment was repeated with 1 µg RNA in order to verify that the single-stranded probe was in excess.

The results of nuclease S1 analysis of modified factor VIII:C mRNA for each construct are shown in Figure 5. Our results indicated that modified factor VIII:C mRNA levels are the same for all three deletions and the full-length factor VIII:C gene. Figure 5A is the analysis for 10 µg of input RNA, and Figure 5B is the analysis for 1 µg of input RNA. Lane 1 in both figures contains as marker 500 cpm of the labeled 477 nucleotide single-stranded DNA fragment used to protect modified factor VIII:C mRNA from S1 digestion; that is, 10% of the input to each hybridization reaction. Lane 2 contained RNA

isolated from BMT10 cells transfected with the signal-minus construct; lane 3, BMT10 cells transfected with the full-length factor VIII:C cDNA (construct 8.1); lane 4: BMT10 cells transfected with modified factor VIII:C cDNA (QD deletion); lane 5: BMT10 cells transfected with modified factor VIII:C cDNA (RE deletion); lane 6: BMT10 cells transfected with the cDNA from the RS deletion; lane 7: marker fragments obtained by digesting pBR322 with HinfI and labeling their 3' ends with [α-^{32}P]dATP and Klenow enzyme (a gift of Richard Tizard). Equal amounts of a protected fragment of the expected length of 300 bases are evident in both figures for the 8.1, QD, RE, and RS constructs. A protected fragment of approximately 220 bases in length for the signal-minus construct is evident in both figures, reflecting the absence of a portion of the DNA sequence encoding the signal peptide.

A comparison of Figures 5A and 5B demonstrates that the input 477 probe is in molar excess during the hybridizations for each construct. Although the modified factor VIII:C activity levels are at least 20-fold higher for the QD and RE deletions compared to the RS and the full-length constructs, the amount of mRNA in all four constructs is very nearly the same. Therefore, the reason for the increase in expression for the QD and RE deletions is post-transcriptional in nature.

I. SOUTHERN ANALYSIS OF PLASMID DNA ISOLATED FROM TRANSFECTED BMT10 CELLS

We conducted this analysis to determine the DNA levels of newly-replicated modified factor VIII:C plasmids for our deletions, in comparison with the full-length gene. Again, this assay assisted in our determination of the reason for the high yields of modified factor VIII:C in our QD and RE deletions.

In order to control for differences in DNA replication in BMT10 cells for the various constructs, we performed a Southern analysis of extrachromosomal DNA isolated from each transfection. We isolated DNA from 100 mm petri dish cultures of BMT10 cells 120 hours after transfection according to the method of Hirt [B. Hirt, J. Mol. Biol., 26, pp. 365-69 (1967)]. For each construction, we digested 0.5 A260 units with DpnI to distinguish newly-replicated (DpnI-resistant) DNA from input methylated bacterial DNA (DpnI-sensitive). We electrophoresed the DNA fragments on a 0.7% agarose gel, and blotted them to GeneScreen Plus to analyze the DNA. The filter was hybridized at 65°C in 1 M NaCl - 50 mM Tris-HCl (pH 7.5) - 0.1% sodium pyrophosphate - 0.2% polyvinyl-pyrrolidone - 0.2% Ficoll - 0.2% BSA - 1% SDS using 10^5 cpm/ml denatured probe. We then washed the filter at 65°C with the same buffer and exposed it overnight at -70°C to Kodak XAR-5 X-ray film backed by a Lightning-Plus intensifying screen (Dupont). The factor VIII:C probe was the 2924 bp EspI fragment isolated from the RE expression plasmid (see Figure 4) and ^{32}P-labeled to a specific activity of 10^9 cpm/µg by the random hexadeoxynucleotide primer method of Feinberg and Vogelstein [A. P. Feinberg and B. Vogelstein, Anal. Biochem., 132, pp. 6-13 (1983)].

Our results, which are depicted in Figure 6, indicate that newly-replicated modified factor VIII:C plasmid DNA levels are the same for all three deletions and the full-length gene. Lane 1 contained the 1 kb ladder obtained from BRL and labeled with T4 DNA polymerase according to the manufacturer's protocol; lane 2: 1 ng supercoiled RE DNA; lane 3: 10 ng supercoiled RE DNA; lane 4: 10 ng RE DNA digested with DpnI; lane 5: DpnI digest of 0.5 A260 units Hirt fraction obtained from BMT10 cells transfected with the signal-minus construct; lane 6: transfected

with the full-length factor VIII:C cDNA (construct 8.1); lane 7: transfected with the QD deletion; lane 8: transfected with the RE deletion: lane 9: transfected with the RS deletion. Figure 6 shows nearly equal amounts of the supercoiled form of each construct after digestion with DpnI (lanes 5-9), thus excluding the possibility that differences in DNA replication enhance the expression of the QD and RE deletions. Lane 2 contains 10^8 molecules of the RE construct and lane 3 contains 10^9 molecules, suggesting that the copy number is approximately 10^3 in the approximately 10^5 cells successfully transfected.

J. CONSTRUCTION OF ARG 740-ASP 1658 (ABBREVIATED RD) DELETION

In this section, we demonstrate how we created the RD deletion which removes the DNA sequence coding on expression from Ser 741 to Ser 1657. We constructed this RD deletion fusion in three steps. In the first step, we digested plasmid QD (Figure 2) with Sau3A between the codons for Ser 1657 and Asp 1658 and between Glu 1794 and Asp 1795. This produced a 411 bp fragment. We also linearized plasmid tsa pML [L. Dailey et al., J. Virol. 54, pp. 739-49 (1985)] at the unique BclI site. We then ligated the 411 base pair fragment derived from plasmid QD with T4 DNA ligase (the ligase for this and the following examples) to the linearized tsa pML at the unique BclI site to generate plasmid 411.BclI, which contains the BclI site on the Asp 1658 side of the 411 bp insert (i.e., 5' to the sequence encoding Asp 1658). Plasmid 411.BclI may be linearized uniquely with BclI, resulting in a 5' GATC overhang which consists of the GAT codon for Asp 1658 and the first base of the CAA codon for Gln 1659.

We also digested plasmid QD with HindIII to cleave the plasmid between Arg 740 and Ser 741 and within the codon for Glu 321 to generate a 1258 bp fragment. We then removed the 5' AGCT overhang with mung bean nuclease and ligated it to the BclI-linearized 411.BclI fragment which had previously been rendered flush by treatment with Klenow enzyme and all four deoxynucleoside triphosphates. This resulted in plasmid RD.411, which contains an Asp718 site 5' to the fusion site within the 1258 bp HindIII fragment. RD.411 contains a PstI site 3' to the fusion site within the 411 bp Sau3A fragment.

Subsequently, we digested plasmid RE (Figure 3B) with Asp718 to cleave within the codon for Trp 585.

We then dephosphorylated the 5' GTAC overhang with calf intestinal phosphatase and then partially digested with PstI. This partial digestion cleaved the linearized RE plasmid between the codons for Ala 1702 and Val 1703, thus removing a 628 bp fragment spanning the RE fusion.

We then cleaved plasmid RD.411 with Asp718 and PstI to generate a 601 bp fragment spanning the RD fusion. We then ligated this fragment to the Asp718-cleaved, PstI-partially cleaved RE plasmid DNA to generate plasmid RD. As demonstrated below, plasmid RD directed the expression of a factor VIII polypeptide with a fusion between Arg 740 and Asp 1658. Cleavage of the RD polypeptide after Arg 740 generates a twochain factor VIII molecule with a mature heavy chain calciumbridged to a δ9 light chain, i.e. a light chain lacking the first 9 amino-terminal amino acids.

K. CONSTRUCTION OF ARG 740-SER 1657 (ABBREVIATED RSD DELETION)

In this section, we demonstrate how we created the RSD deletion which removes the DNA

sequence coding on expression for Ser 741 to Gln 1656 of the mature polypeptide. Initially, we constructed plasmid 411.BclI and linearized it with BclI as described in example "J". Subsequently, we digested plasmid QD with HindIII, cleaving the plasmid between the codons for Arg 740 and Ser 741 and within the codon for Glu 321 to generate a 1258 bp fragment. We preserved the AGC codon within the 5' AGCT overhang with Klenow enzyme and dATP, dGTP and dCTP and then removed the leftover 5' T overhang with mung bean nuclease. We then ligated this modified HindIII fragment to BclI-linearized 411.BclI, which had been previously treated with Klenow enzyme and all four deoxynucleoside triphosphates, to produce plasmid RSD.411, which contains an Asp718 site 5' to the fusion site within the 1258 bp HindIII fragment and a PstI site 3' to the fusion site within the 411 bp Sau3A fragment.

We then prepared Asp718-cleaved, PstI-partially cleaved RE plasmid DNA as described in example "D". Subsequently, we cleaved plasmid RSD.411 with Asp718 and PstI and ligated the resulting 604 bp fragment spanning the RSD fusion to the Asp718-cleaved, PstI-partially cleaved RE plasmid DNA to generate plasmid RSD. Upon expression, the RSD plasmid encoded a factor VIII polypeptide with a fusion between Arg 740 and Ser 1657. A cleavage of RSD polypeptide after Arg 740 generates a 2-chain factor VIII molecule with a mature heavy chain and a delta 8 light chain, i.e. a light chain lacking the first eight amino terminal amino acids. Furthermore, because in the primary translation product Ser is also at position 741, RSD may also be viewed as a fusion between Ser 741 and Asp 1658. A cleavage after Ser 741 may generate a 2-chain factor VIII molecule with a heavy chain terminating at Ser 741 and a δ9 light chain.

L. TRANSFECTION OF AFRICAN GREEN MONKEY KIDNEY CELLS

We first produced African green monkey kidney cell line 6L by cotransfecting cell line BSC40 (BSC1 African green monkey kidney cells which have been adapted to grow at 40°C), [W. Brackman and D. Nathan, Proc. Natl. Acad. Sci. USA, 71, pp. 942-46 (1974)] with pLTRtsA58 and with pY3, which has a transcription unit for hygromycin B phosphotranferase [K. Blochlinger, and A. Diggelmann, Mol. Cell Biol. 4, p. 2929-31 (1984)]. Plasmid LTRtsA58 contains a transcription unit for a temperature sensitive SV40 T-antigen allele. A mutant tsA58 virus is a temperature-sensitive mutant of SV40 which does not produce progeny at 39°C. The large T-antigen protein specified by the tsA58 mutant is much more labile at the nonpermissive temperature than wild type large T-antigen protein [H. Tegtmeyer et al., J. Virol 16, pp. 168-78 (1975). The resulting cell line 6L inducibly expresses SV40 T-antigen at 33°C.

We then transfected 6L cells with super-coiled expression plasmids RD or RSD. The transfection was carried out using the DEAE-dextran technique and chloroquine as described in Example "F". We then incubated the transfected cells at 33°C. During incubation, the transfected cells synthesized and secreted modified factor VIII:C into the culture fluid. The transfectants will secrete modified factor VIII:C for up to 120 hours. For most assays, the cm^2/ml ratio was approximately 5.5; that is, a confluent monolayer of 6L transfectants in a 100mm Petri dish ($55cm^2$) was covered with 10ml culture fluid.

M. FACTOR VIII:C ACTIVITY ASSAY

We assayed the RE (Example D), RD and RSD expression constructs for factor VIII:C production after transfection and incubation at 33°C for three

days using KabiVitrum's Coatest® factor VIII assay kit adapted to a 96 well plate. Cells transfected with plasmid RE produced culture fluid having a factor VIII concentration which was 0.48% plasma level [normal plasma factor VIII concentration is approximately 150 ng/ml]. Cells transfected with plasmid RD produced culture fluid having a factor VIII concentration which was 0.41% plasma level. Cells transfected with plasmid RSD produced culture fluid having a factor VIII concentration which was 0.71% plasma level.

In a similar assay, cells transfected with plasmids RE or RSD which had been incubated at 33°C for three days and then for an additional two days, yielded the following factor VIII concentrations in the cell culture fluid:

Factor VIII:C Concentration In Culture Fluid As % Of Plasma Level

	3 Days	5 Days
RE Transfected Cells	0.30%	0.77%
RSD Transfected Cells	1.50%	1.16%

Microorganisms, recombinant DNA molecules and the modified factor VIII:C DNA coding sequences of this invention are exemplified by a culture deposited in the culture collection of the American Type Culture Collection, in Rockville, Maryland, on July 22, 1986, and identified there as:

E.coli HB101 (RE)

This culture was assigned ATCC accession number 53517. Two additional cultures were deposited in the American

Type Culture Collection, in Rockville, Maryland on July 27, 1987, and identified there as:

Ad.RD.2 [E.coli HB101 (RD)], having ATCC accession number 67475; and Ad.RSD.1.2 [E.coli HB101 (RSD)], having ATCC accession number 67476.

While we have hereinbefore presented a number of embodiments of this invention, it is apparent that our basic construction can be altered to provide other embodiments which utilize the processes and compositions of this invention. Therefore, it will be appreciated that the scope of this invention is to be defined by the claims appended hereto rather than by the specific embodiments which have been presented hereinbefore by way of example.

We claim:

1. A recombinant DNA molecule characterized by a DNA sequence coding on expression for a modified factor VIII:C-like polypeptide, said DNA sequence containing a deletion of a major part of the DNA sequence which codes on expression for the maturation polypeptide of factor VIII:C.

2. The recombinant DNA molecule according to claim 1, wherein the deletion is all of the DNA sequence which codes on expression for the maturation polypeptide of factor VIII:C.

3. The recombinant DNA molecule according to claim 1, wherein the DNA sequence coding on expression for the modified factor VIII:C-like polypeptide is selected from the group consisting of:

ATG GCC ACC AGA AGA TAC TAC CTG GGT GCA GTG GAA CTG
TCA TGG GAC TAT ATG CAA AGT GAT CTC GGT GAG CTG CCT
GTG GAC GCA AGA TTT CCT CCT AGA GTG CCA AAA TCT TTT
CCA TTC AAC ACC TCA GTC GTG TAC AAA AAG ACT CTG TTT
GTA GAA TTC ACG GAT CAC CTT TTC AAC ATC GCT AAG CCA
AGG CCA CCC TGG ATG GGT CTG CTA GGT CCT ACC ATC CAG
GCT GAG GTT TAT GAT ACA GTG GTC ATT ACA CTT AAG AAC
ATG GCT TCC CAT CCT GTC AGT CTT CAT GCT GTT GGT GTA
TCC TAC TGG AAA GCT TCT GAG GGA GCT GAA TAT GAT GAT
CAG ACC AGT CAA AGG GAG AAA GAA GAT GAT AAA GTC TTC
CCT GGT GGA AGC CAT ACA TAT GTC TGG CAG GTC CTG AAA
GAG AAT GGT CCA ATG GCC TCT GAC CCA CTG TGC CTT ACC
TAC TCA TAT CTT TCT CAT GTG GAC CTG GTA AAA GAC TTG
AAT TCA GGC CTC ATT GGA GCC CTA CTA GTA TGT AGA GAA
GGG AGT CTG GCC AAG GAA AAG ACA CAC ACC TTG CAC AAA
TTT ATA CTA CTT TTT GCT GTA TTT GAT GAA GGG AAA AGT
TGG CAC TCA GAA ACA AAG AAC TCC TTG ATG CAG GAT AGG
GAT GCT GCA TCT GCT CGG GCC TGG CCT AAA ATG CAC ACA
GTC AAT GGT TAT GTA AAC AGG TCT CTG(CTA) CCA GGT CTG
ATT GGA TGC CAC AGG AAA TCA GTC TAT TGG CAT GTG ATT

GGA ATG GGC ACC ACT CCT GAA GTG CAC TCA ATA TTC CTC
GAA GGT CAC ACA TTT CTT GTG AGG AAC CAT CGC CAG GCG
TCC TTG GAA ATC TCG CCA ATA ACT TTC CTT ACT GCT CAA
ACA CTC TTG ATG GAC CTT GGA CAG TTT CTA CTG TTT TGT
CAT ATC TCT TCC CAC CAA CAT GAT GGC ATG GAA GCT TAT
GTC AAA GTA GAC AGC TGT CCA GAG GAA CCC CAA CTA CGA
ATG AAA AAT AAT GAA GAA GCG GAA GAC TAT GAT GAT GAT
CTT ACT GAT TCT GAA ATG GAT GTG GTC AGG TTT GAT GAT
GAC AAC TCT CCT TCC TTT ATC CAA ATT CGC TCA GTT GCC
AAG AAG CAT CCT AAA ACT TGG GTA CAT TAC ATT GCT GCT
GAA GAG GAG GAC TGG GAC TAT GCT CCC TTA GTC CTC GCC
CCC GAT GAC AGA AGT TAT AAA AGT CAA TAT TTG AAC AAT
GGC CCT CAG CGG ATT GGT AGG AAG TAC AAA AAA GTC CGA
TTT ATG GCA TAC ACA GAT GAA ACC TTT AAG ACT CGT GAA
GCT ATT CAG CAT GAA TCA GGA ATC TTG GGA CCT TTA CTT
TAT GGG GAA GTT GGA GAC ACA CTG TTG ATT ATA TTT AAG
AAT CAA GCA AGC AGA CCA TAT AAC ATC TAC CCT CAC GGA
ATC ACT GAT GTC CGT CCT TTG TAT TCA AGG AGA TTA CCA
AAA GGT GTA AAA CAT TTG AAG GAT TTT CCA ATT CTG CCA
GGA GAA ATA TTC AAA TAT AAA TGG ACA GTG ACT GTA GAA
GAT GGG CCA ACT AAA TCA GAT CCT CGG TGC CTG ACC CGC
TAT TAC TCT AGT TTC GTT AAT ATG GAG AGA GAT CTA GCT
TCA GGA CTC ATT GGC CCT CTC CTC ATC TGC TAC AAA GAA
TCT GTA GAT CAA AGA GGA AAC CAG ATA ATG TCA GAC AAG
AGG AAT GTC ATC CTG TTT TCT GTA TTT GAT GAG AAC CGA
AGC TGG TAC CTC ACA GAG AAT ATA CAA CGC TTT CTC CCC
AAT CCA GCT GGA GTG CAG CTT GAG GAT CCA GAG TTC CAA
GCC TCC AAC ATC ATG CAC AGC ATC AAT GGC TAT GTT TTT
GAT AGT TTG CAG TTG TCA GTT TGT TTG CAT GAG GTG GCA
TAC TGG TAC ATT CTA AGC ATT GGA GCA CAG ACT GAC TTC
CTT TCT GTC TTC TTC TCT GGA TAT ACC TTC AAA CAC AAA
ATG GTC TAT GAA GAC ACA CTC ACC CTA TTC CCA TTC TCA
GGA GAA ACT GTC TTC ATG TCG ATG GAA AAC CCA GGT CTA
TGG ATT CTG GGG TGC CAC AAC TCA GAC TTT CGG AAC AGA
GGC ATG ACC GCC TTA CTG AAG GTT TCT AGT TGT GAC AAG
AAC ACT GGT GAT TAT TAC GAG GAC AGT TAT GAA GAT ATT
TCA GCA TAC TTG CTG AGT AAA AAC AAT GCC ATT GAA CCA

AGA AGC TTC TCC CAG GAT CCT CTT GCT TGG GAT AAC CAC
TAT GGT ACT CAG ATA CCA AAA GAA GAG TGG AAA TCC CAA
GAG AAG TCA CCA GAA AAA ACA GCT TTT AAG AAA AAG GAT
ACC ATT TTG TCC CTG AAC GCT TGT GAA AGC AAT CAT GCA
ATA GCA GCA ATA AAT GAG GGA CAA AAT AAG CCC GAA ATA
GAA GTC ACC TGG GCA AAG CAA GGT AGG ACT GAA AGG CTG
TGC TCT CAA AAC CCA CCA GTC TTG AAA CGC CAT CAA CGG
GAA ATA ACT CGT ACT ACT CTT CAG TCA GAT CAA GAG GAA
ATT GAC TAT GAT GAT ACC ATA TCA GTT GAA ATG AAG AAG
GAA GAT TTT GAC ATT TAT GAT GAG GAT GAA AAT CAG AGC
CCC CGC AGC TTT CAA AAG AAA ACA CGA CAC TAT TTT ATT
GCT GCA GTG GAG AGG CTC TGG GAT TAT GGG ATG AGT AGC
TCC CCA CAT GTT CTA AGA AAC AGG GCT CAG AGT GGC AGT
GTC CCT CAG TTC AAG AAA GTT GTT TTC CAG GAA TTT ACT
GAT GGC TCC TTT ACT CAG CCC TTA TAC CGT GGA GAA CTA
AAT GAA CAT TTG GGA CTC CTG GGG CCA TAT ATA AGA GCA
GAA GTT GAA GAT AAT ATC ATG GTA ACT TTC AGA AAT CAG
GCC TCT CGT CCC TAT TCC TTC TAT TCT AGC CTT ATT TCT
TAT GAG GAA GAT CAG AGG CAA GGA GCA GAA CCT AGA AAA
AAC TTT GTC AAG CCT AAT GAA ACC AAA ACT TAC TTT TGG
AAA GTG CAA CAT CAT ATG GCA CCC ACT AAA GAT GAG TTT
GAC TGC AAA GCC TGG GCT TAT TTC TCT GAT GTT GAC CTG
GAA AAA GAT GTG CAC TCA GGC CTG ATT GGA CCC CTT CTG
GTC TGC CAC ACT AAC ACA CTG AAC CCT GCT CAT GGG AGA
CAA GTG ACA GTA CAG GAA TTT GCT CTG TTT TTC(CTC) ACC
ATC TTT GAT GAG ACC AAA AGC TGG TAC TTC ACT GAA AAT
ATG GAA AGA AAC TGC AGG GCT CCC TGC AAT ATC CAG ATG
GAA GAT CCC ACT TTT AAA GAG AAT TAT CGC TTC CAT GCA
ATC AAT GGC TAC ATA ATG GAT ACA CTA CCT GGC TTA GTA
ATG GCT CAG GAT CAA AGG ATT CGA TGG TAT CTG CTC AGC
ATG GGC AGC AAT GAA AAC ATC CAT TCT ATT CAT TTC AGT
GGA CAT GTG TTC ACT GTA CGA AAA AAA GAG GAG TAT AAA
ATG GCA CTG TAC AAT CTC TAT CCA GGT GTT TTT GAG ACA
GTG GAA ATG TTA CCA TCC AAA GCT GGA ATT TGG CGG GTG
GAA TGC CTT ATT GGC GAG CAT CTA CAT GCT GGG ATG AGC
ACA CTT TTT CTG GTG TAC AGC AAT AAG TGT CAG ACT CCC
CTG GGA ATG GCT TCT GGA CAC ATT AGA GAT TTT CAG ATT

ACA GCT TCA GGA CAA TAT GGA CAG TGG GCC CCA AAG CTG
GCC AGA CTT CAT TAT TCC GGA TCA ATC AAT GCC TGG AGC
ACC AAG GAG CCC TTT TCT TGG ATC AAG GTG GAT CTG TTG
GCA CCA ATG ATT ATT CAC GGC ATC AAG ACC CAG GGT GCC
CGT CAG AAG TTC TCC AGC CTC TAC ATC TCT CAG TTT ATC
ATC ATG TAT AGT CTT GAT GGG AAG AAG TGG CAG ACT TAT
CGA GGA AAT TCC ACT GGA ACC TTA ATG GTC TTC TTT GGC
AAT GTG GAT TCA TCT GGG ATA AAA CAC AAT ATT TTT AAC
CCT CCA ATT ATT GCT CGA TAC ATC CGT TTG CAC CCA ACT
CAT TAT AGC ATT CGC AGC ACT CTT CGC ATG GAG TTG ATG
GGC TGT GAT TTA AAT AGT TGC AGC ATG CCA TTG GGA ATG
GAG AGT AAA GCA ATA TCA GAT GCA CAG ATT ACT GCT TCA
TCC TAC TTT ACC AAT ATG TTT GCC ACC TGG TCT CCT TCA
AAA GCT CGA CTT CAC CTC CAA GGG AGG AGT AAT GCC TGG
AGA CCT CAG GTG AAT AAT CCA AAA GAG TGG CTG CAA GTG
GAC TTC CAG AAG ACA ATG AAA GTC ACA GGA GTA ACT ACT
CAG GGA GTA AAA TCT CTG CTT ACC AGC ATG TAT GTG AAG
GAG TTC CTC ATC TCC AGC AGT CAA GAT GGC CAT CAG TGG
ACT CTC TTT TTT CAG AAT GGC AAA GTA AAG GTT TTT CAG
GGA AAT CAA GAC TCC TTC ACA CCT GTG GTG AAC TCT CTA
GAC CCA CCG TTA CTG ACT CGC TAC CTT CGA ATT CAC CCC
CAG AGT TGG GTG CAC CAG ATT GCC CTG AGG ATG GAG GTT
CTG GGC TGC GAG GCA CAG GAC CTC TAC;
GCC ACC AGA AGA TAC TAC CTG GGT GCA GTG GAA CTG TCA
TGG GAC TAT ATG CAA AGT GAT CTC GGT GAG CTG CCT GTG
GAC GCA AGA TTT CCT CCT AGA GTG CCA AAA TCT TTT CCA
TTC AAC ACC TCA GTC GTG TAC AAA AAG ACT CTG TTT GTA
GAA TTC ACG GAT CAC CTT TTC AAC ATC GCT AAG CCA AGG
CCA CCC TGG ATG GGT CTG CTA GGT CCT ACC ATC CAG GCT
GAG GTT TAT GAT ACA GTG GTC ATT ACA CTT AAG AAC ATG
GCT TCC CAT CCT GTC AGT CTT CAT GCT GTT GGT GTA TCC
TAC TGG AAA GCT TCT GAG GGA GCT GAA TAT GAT GAT CAG
ACC AGT CAA AGG GAG AAA GAA GAT GAT AAA GTC TTC CCT
GGT GGA AGC CAT ACA TAT GTC TGG CAG GTC CTG AAA GAG
AAT GGT CCA ATG GCC TCT GAC CCA CTG TGC CTT ACC TAC
TCA TAT CTT TCT CAT GTG GAC CTG GTA AAA GAC TTG AAT
TCA GGC CTC ATT GGA GCC CTA CTA GTA TGT AGA GAA GGG

AGT CTG GCC AAG GAA AAG ACA CAC ACC TTG CAC AAA TTT
ATA CTA CTT TTT GCT GTA TTT GAT GAA GGG AAA AGT TGG
CAC TCA GAA ACA AAG AAC TCC TTG ATG CAG GAT AGG GAT
GCT GCA TCT GCT CGG GCC TGG CCT AAA ATG CAC ACA GTC
AAT GGT TAT GTA AAC AGG TCT CTG(CTA) CCA GGT CTG ATT
GGA TGC CAC AGG AAA TCA GTC TAT TGG CAT GTG ATT GGA
ATG GGC ACC ACT CCT GAA GTG CAC TCA ATA TTC CTC GAA
GGT CAC ACA TTT CTT GTG AGG AAC CAT CGC CAG GCG TCC
TTG GAA ATC TCG CCA ATA ACT TTC CTT ACT GCT CAA ACA
CTC TTG ATG GAC CTT GGA CAG TTT CTA CTG TTT TGT CAT
ATC TCT TCC CAC CAA CAT GAT GGC ATG GAA GCT TAT GTC
AAA GTA GAC AGC TGT CCA GAG GAA CCC CAA CTA CGA ATG
AAA AAT AAT GAA GAA GCG GAA GAC TAT GAT GAT GAT CTT
ACT GAT TCT GAA ATG GAT GTG GTC AGG TTT GAT GAT GAC
AAC TCT CCT TCC TTT ATC CAA ATT CGC TCA GTT GCC AAG
AAG CAT CCT AAA ACT TGG GTA CAT TAC ATT GCT GCT GAA
GAG GAG GAC TGG GAC TAT GCT CCC TTA GTC CTC GCC CCC
GAT GAC AGA AGT TAT AAA AGT CAA TAT TTG AAC AAT GC
CCT CAG CGG ATT GGT AGG AAG TAC AAA AAA GTC CGA TTT
ATG GCA TAC ACA GAT GAA ACC TTT AAG ACT CGT GAA GCT
ATT CAG CAT GAA TCA GGA ATC TTG GGA CCT TTA CTT TAT
GGG GAA GTT GGA GAC ACA CTG TTG ATT ATA TTT AAG AAT
CAA GCA AGC AGA CCA TAT AAC ATC TAC CCT CAC GGA ATC
ACT GAT GTC CGT CCT TTG TAT TCA AGG AGA TTA CCA AAA
GGT GTA AAA CAT TTG AAG GAT TTT CCA ATT CTG CCA GGA
GAA ATA TTC AAA TAT AAA TGG ACA GTG ACT GTA GAA GAT
GGG CCA ACT AAA TCA GAT CCT CGG TGC CTG ACC CGC TAT
TAC TCT AGT TTC GTT AAT ATG GAG AGA GAT CTA GCT TCA
GGA CTC ATT GGC CCT CTC CTC ATC TGC TAC AAA GAA TCT
GTA GAT CAA AGA GGA AAC CAG ATA ATG TCA GAC AAG AGG
AAT GTC ATC CTG TTT TCT GTA TTT GAT GAG AAC CGA AGC
TGG TAC CTC ACA GAG AAT ATA CAA CGC TTT CTC CCC AAT
CCA GCT GGA GTG CAG CTT GAG GAT CCA GAG TTC CAA GCC
TCC AAC ATC ATG CAC AGC ATC AAT GGC TAT GTT TTT GAT
AGT TTG CAG TTG TCA GTT TGT TTG CAT GAG GTG GCA TAC
TGG TAC ATT CTA AGC ATT GGA GCA CAG ACT GAC TTC CTT
TCT GTC TTC TTC TCT GGA TAT ACC TTC AAA CAC AAA ATG

GTC TAT GAA GAC ACA CTC ACC CTA TTC CCA TTC TCA GGA
GAA ACT GTC TTC ATG TCG ATG GAA AAC CCA GGT CTA TGG
ATT CTG GGG TGC CAC AAC TCA GAC TTT CGG AAC AGA GGC
ATG ACC GCC TTA CTG AAG GTT TCT AGT TGT GAC AAG AAC
ACT GGT GAT TAT TAC GAG GAC AGT TAT GAA GAT ATT TCA
GCA TAC TTG CTG AGT AAA AAC AAT GCC ATT GAA CCA AGA
AGC TTC TCC CAG GAT CCT CTT GCT TGG GAT AAC CAC TAT
GGT ACT CAG ATA CCA AAA GAA GAG TGG AAA TCC CAA GAG
AAG TCA CCA GAA AAA ACA GCT TTT AAG AAA AAG GAT ACC
ATT TTG TCC CTG AAC GCT TGT GAA AGC AAT CAT GCA ATA
GCA GCA ATA AAT GAG GGA CAA AAT AAG CCC GAA ATA GAA
GTC ACC TGG GCA AAG CAA GGT AGG ACT GAA AGG CTG TGC
TCT CAA AAC CCA CCA GTC TTG AAA CGC CAT CAA CGG GAA
ATA ACT CGT ACT ACT CTT CAG TCA GAT CAA GAG GAA ATT
GAC TAT GAT GAT ACC ATA TCA GTT GAA ATG AAG AAG GAA
GAT TTT GAC ATT TAT GAT GAG GAT GAA AAT CAG AGC CCC
CGC AGC TTT CAA AAG AAA ACA CGA CAC TAT TTT ATT GCT
GCA GTG GAG AGG CTC TGG GAT TAT GGG ATG AGT AGC TCC
CCA CAT GTT CTA AGA AAC AGG GCT CAG AGT GGC AGT GTC
CCT CAG TTC AAG AAA GTT GTT TTC CAG GAA TTT ACT GAT
GGC TCC TTT ACT CAG CCC TTA TAC CGT GGA GAA CTA AAT
GAA CAT TTG GGA CTC CTG GGG CCA TAT ATA AGA GCA GAA
GTT GAA GAT AAT ATC ATG GTA ACT TTC AGA AAT CAG GCC
TCT CGT CCC TAT TCC TTC TAT TCT AGC CTT ATT TCT TAT
GAG GAA GAT CAG AGG CAA GGA GCA GAA CCT AGA AAA AAC
TTT GTC AAG CCT AAT GAA ACC AAA ACT TAC TTT TGG AAA
GTG CAA CAT CAT ATG GCA CCC ACT AAA GAT GAG TTT GAC
TGC AAA GCC TGG GCT TAT TTC TCT GAT GTT GAC CTG GAA
AAA GAT GTG CAC TCA GGC CTG ATT GGA CCC CTT CTG GTC
TGC CAC ACT AAC ACA CTG AAC CCT GCT CAT GGG AGA CAA
GTG ACA GTA CAG GAA TTT GCT CTG TTT TTC(CTC) ACC ATC
TTT GAT GAG ACC AAA AGC TGG TAC TTC ACT GAA AAT ATG
GAA AGA AAC TGC AGG GCT CCC TGC AAT ATC CAG ATG GAA
GAT CCC ACT TTT AAA GAG AAT TAT CGC TTC CAT GCA ATC
AAT GGC TAC ATA ATG GAT ACA CTA CCT GGC TTA GTA ATG
GCT CAG GAT CAA AGG ATT CGA TGG TAT CTG CTC AGC ATG
GGC AGC AAT GAA AAC ATC CAT TCT ATT CAT TTC AGT GGA

CAT GTG TTC ACT GTA CGA AAA AAA GAG GAG TAT AAA ATG
GCA CTG TAC AAT CTC TAT CCA GGT GTT TTT GAG ACA GTG
GAA ATG TTA CCA TCC AAA GCT GGA ATT TGG CGG GTG GAA
TGC CTT ATT GGC GAG CAT CTA CAT GCT GGG ATG AGC ACA
CTT TTT CTG GTG TAC AGC AAT AAG TGT CAG ACT CCC CTG
GGA ATG GCT TCT GGA CAC ATT AGA GAT TTT CAG ATT ACA
GCT TCA GGA CAA TAT GGA CAG TGG GCC CCA AAG CTG GCC
AGA CTT CAT TAT TCC GGA TCA ATC AAT GCC TGG AGC ACC
AAG GAG CCC TTT TCT TGG ATC AAG GTG GAT CTG TTG GCA
CCA ATG ATT ATT CAC GGC ATC AAG ACC CAG GGT GCC CGT
CAG AAG TTC TCC AGC CTC TAC ATC TCT CAG TTT ATC ATC
ATG TAT AGT CTT GAT GGG AAG AAG TGG CAG ACT TAT CGA
GGA AAT TCC ACT GGA ACC TTA ATG GTC TTC TTT GGC AAT
GTG GAT TCA TCT GGG ATA AAA CAC AAT ATT TTT AAC CCT
CCA ATT ATT GCT CGA TAC ATC CGT TTG CAC CCA ACT CAT
TAT AGC ATT CGC AGC ACT CTT CGC ATG GAG TTG ATG GGC
TGT GAT TTA AAT AGT TGC AGC ATG CCA TTG GGA ATG GAG
AGT AAA GCA ATA TCA GAT GCA CAG ATT ACT GCT TCA TCC
TAC TTT ACC AAT ATG TTT GCC ACC TGG TCT CCT TCA AAA
GCT CGA CTT CAC CTC CAA GGG AGG AGT AAT GCC TGG AGA
CCT CAG GTG AAT AAT CCA AAA GAG TGG CTG CAA GTG GAC
TTC CAG AAG ACA ATG AAA GTC ACA GGA GTA ACT ACT CAG
GGA GTA AAA TCT CTG CTT ACC AGC ATG TAT GTG AAG GAG
TTC CTC ATC TCC AGC AGT CAA GAT GGC CAT CAG TGG ACT
CTC TTT TTT CAG AAT GGC AAA GTA AAG GTT TTT CAG GGA
AAT CAA GAC TCC TTC ACA CCT GTG GTG AAC TCT CTA GAC
CCA CCG TTA CTG ACT CGC TAC CTT CGA ATT CAC CCC CAG
AGT TGG GTG CAC CAG ATT GCC CTG AGG ATG GAG GTT CTG
GGC TGC GAG GCA CAG GAC CTC TAC;
ATG GCC ACC AGA AGA TAC TAC CTG GGT GCA GTG GAA CTG
TCA TGG GAC TAT ATG CAA AGT GAT CTC GGT GAG CTG CCT
GTG GAC GCA AGA TTT CCT CCT AGA GTG CCA AAA TCT TTT
CCA TTC AAC ACC TCA GTC GTG TAC AAA AAG ACT CTG TTT
GTA GAA TTC ACG GAT CAC CTT TTC AAC ATC GCT AAG CCA
AGG CCA CCC TGG ATG GGT CTG CTA GGT CCT ACC ATC CAG
GCT GAG GTT TAT GAT ACA GTG GTC ATT ACA CTT AAG AAC
ATG GCT TCC CAT CCT GTC AGT CTT CAT GCT GTT GGT GTA

TCC TAC TGG AAA GCT TCT GAG GGA GCT GAA TAT GAT GAT
CAG ACC AGT CAA AGG GAG AAA GAA GAT GAT AAA GTC TTC
CCT GGT GGA AGC CAT ACA TAT GTC TGG CAG GTC CTG AAA
GAG AAT GGT CCA ATG GCC TCT GAC CCA CTG TGC CTT ACC
TAC TCA TAT CTT TCT CAT GTG GAC CTG GTA AAA GAC TTG
AAT TCA GGC CTC ATT GGA GCC CTA CTA GTA TGT AGA GAA
GGG AGT CTG GCC AAG GAA AAG ACA CAC ACC TTG CAC AAA
TTT ATA CTA CTT TTT GCT GTA TTT GAT GAA GGG AAA AGT
TGG CAC TCA GAA ACA AAG AAC TCC TTG ATG CAG GAT AGG
GAT GCT GCA TCT GCT CGG GCC TGG CCT AAA ATG CAC ACA
GTC AAT GGT TAT GTA AAC AGG TCT CTG(CTA) CCA GGT CTG
ATT GGA TGC CAC AGG AAA TCA GTC TAT TGG CAT GTG ATT
GGA ATG GGC ACC ACT CCT GAA GTG CAC TCA ATA TTC CTC
GAA GGT CAC ACA TTT CTT GTG AGG AAC CAT CGC CAG GCG
TCC TTG GAA ATC TCG CCA ATA ACT TTC CTT ACT GCT CAA
ACA CTC TTG ATG GAC CTT GGA CAG TTT CTA CTG TTT TGT
CAT ATC TCT TCC CAC CAA CAT GAT GGC ATG GAA GCT TAT
GTC AAA GTA GAC AGC TGT CCA GAG GAA CCC CAA CTA CGA
ATG AAA AAT AAT GAA GAA GCG GAA GAC TAT GAT GAT GAT
CTT ACT GAT TCT GAA ATG GAT GTG GTC AGG TTT GAT GAT
GAC AAC TCT CCT TCC TTT ATC CAA ATT CGC TCA GTT GCC
AAG AAG CAT CCT AAA ACT TGG GTA CAT TAC ATT GCT GCT
GAA GAG GAG GAC TGG GAC TAT GCT CCC TTA GTC CTC GCC
CCC GAT GAC AGA AGT TAT AAA AGT CAA TAT TTG AAC AAT
GGC CCT CAG CGG ATT GGT AGG AAG TAC AAA AAA GTC CGA
TTT ATG GCA TAC ACA GAT GAA ACC TTT AAG ACT CGT GAA
GCT ATT CAG CAT GAA TCA GGA ATC TTG GGA CCT TTA CTT
TAT GGG GAA GTT GGA GAC ACA CTG TTG ATT ATA TTT AAG
AAT CAA GCA AGC AGA CCA TAT AAC ATC TAC CCT CAC GGA
ATC ACT GAT GTC CGT CCT TTG TAT TCA AGG AGA TTA CCA
AAA GGT GTA AAA CAT TTG AAG GAT TTT CCA ATT CTG CCA
GGA GAA ATA TTC AAA TAT AAA TGG ACA GTG ACT GTA GAA
GAT GGG CCA ACT AAA TCA GAT CCT CGG TGC CTG ACC CGC
TAT TAC TCT AGT TTC GTT AAT ATG GAG AGA GAT CTA GCT
TCA GGA CTC ATT GGC CCT CTC CTC ATC TGC TAC AAA GAA
TCT GTA GAT CAA AGA GGA AAC CAG ATA ATG TCA GAC AAG
AGG AAT GTC ATC CTG TTT TCT GTA TTT GAT GAG AAC CGA

AGC TGG TAC CTC ACA GAG AAT ATA CAA CGC TTT CTC CCC
AAT CCA GCT GGA GTG CAG CTT GAG GAT CCA GAG TTC CAA
GCC TCC AAC ATC ATG CAC AGC ATC AAT GGC TAT GTT TTT
GAT AGT TTG CAG TTG TCA GTT TGT TTG CAT GAG GTG GCA
TAC TGG TAC ATT CTA AGC ATT GGA GCA CAG ACT GAC TTC
CTT TCT GTC TTC TTC TCT GGA TAT ACC TTC AAA CAC AAA
ATG GTC TAT GAA GAC ACA CTC ACC CTA TTC CCA TTC TCA
GGA GAA ACT GTC TTC ATG TCG ATG GAA AAC CCA GGT CTA
TGG ATT CTG GGG TGC CAC AAC TCA GAC TTT CGG AAC AGA
GGC ATG ACC GCC TTA CTG AAG GTT TCT AGT TGT GAC AAG
AAC ACT GGT GAT TAT TAC GAG GAC AGT TAT GAA GAT ATT
TCA GCA TAC TTG CTG AGT AAA AAC AAT GCC ATT GAA CCA
AGA GAA ATA ACT CGT ACT ACT CTT CAG TCA GAT CAA GAG
GAA ATT GAC TAT GAT GAT ACC ATA TCA GTT GAA ATG AAG
AAG GAA GAT TTT GAC ATT TAT GAT GAG GAT GAA AAT CAG
AGC CCC CGC AGC TTT CAA AAG AAA ACA CGA CAC TAT TTT
ATT GCT GCA GTG GAG AGG CTC TGG GAT TAT GGG ATG AGT
AGC TCC CCA CAT GTT CTA AGA AAC AGG GCT CAG AGT GGC
AGT GTC CCT CAG TTC AAG AAA GTT GTT TTC CAG GAA TTT
ACT GAT GGC TCC TTT ACT CAG CCC TTA TAC CGT GGA GAA
CTA AAT GAA CAT TTG GGA CTC CTG GGG CCA TAT ATA AGA
GCA GAA GTT GAA GAT AAT ATC ATG GTA ACT TTC AGA AAT
CAG GCC TCT CGT CCC TAT TCC TTC TAT TCT AGC CTT ATT
TCT TAT GAG GAA GAT CAG AGG CAA GGA GCA GAA CCT AGA
AAA AAC TTT GTC AAG CCT AAT GAA ACC AAA ACT TAC TTT
TGG AAA GTG CAA CAT CAT ATG GCA CCC ACT AAA GAT GAG
TTT GAC TGC AAA GCC TGG GCT TAT TTC TCT GAT GTT GAC
CTG GAA AAA GAT GTG CAC TCA GGC CTG ATT GGA CCC CTT
CTG GTC TGC CAC ACT AAC ACA CTG AAC CCT GCT CAT GGG
AGA CAA GTG ACA GTA CAG GAA TTT GCT CTG TTT TTC(CTC)
ACC ATC TTT GAT GAG ACC AAA AGC TGG TAC TTC ACT GAA
AAT ATG GAA AGA AAC TGC AGG GCT CCC TGC AAT ATC CAG
ATG GAA GAT CCC ACT TTT AAA GAG AAT TAT CGC TTC CAT
GCA ATC AAT GGC TAC ATA ATG GAT ACA CTA CCT GGC TTA
GTA ATG GCT CAG GAT CAA AGG ATT CGA TGG TAT CTG CTC
AGC ATG GGC AGC AAT GAA AAC ATC CAT TCT ATT CAT TTC
AGT GGA CAT GTG TTC ACT GTA CGA AAA AAA GAG GAG TAT

AAA ATG GCA CTG TAC AAT CTC TAT CCA GGT GTT TTT GAG
ACA GTG GAA ATG TTA CCA TCC AAA GCT GGA ATT TGG CGG
GTG GAA TGC CTT ATT GGC GAG CAT CTA CAT GCT GGG ATG
AGC ACA CTT TTT CTG GTG TAC AGC AAT AAG TGT CAG ACT
CCC CTG GGA ATG GCT TCT GGA CAC ATT AGA GAT TTT CAG
ATT ACA GCT TCA GGA CAA TAT GGA CAG TGG GCC CCA AAG
CTG GCC AGA CTT CAT TAT TCC GGA TCA ATC AAT GCC TGG
AGC ACC AAG GAG CCC TTT TCT TGG ATC AAG GTG GAT CTG
TTG GCA CCA ATG ATT ATT CAC GGC ATC AAG ACC CAG GGT
GCC CGT CAG AAG TTC TCC AGC CTC TAC ATC TCT CAG TTT
ATC ATC ATG TAT AGT CTT GAT GGG AAG AAG TGG CAG ACT
TAT CGA GGA AAT TCC ACT GGA ACC TTA ATG GTC TTC TTT
GGC AAT GTG GAT TCA TCT GGG ATA AAA CAC AAT ATT TTT
AAC CCT CCA ATT ATT GCT CGA TAC ATC CGT TTG CAC CCA
ACT CAT TAT AGC ATT CGC AGC ACT CTT CGC ATG GAG TTG
ATG GGC TGT GAT TTA AAT AGT TGC AGC ATG CCA TTG GGA
ATG GAG AGT AAA GCA ATA TCA GAT GCA CAG ATT ACT GCT
TCA TCC TAC TTT ACC AAT ATG TTT GCC ACC TGG TCT CCT
TCA AAA GCT CGA CTT CAC CTC CAA GGG AGG AGT AAT GCC
TGG AGA CCT CAG GTG AAT AAT CCA AAA GAG TGG CTG CAA
GTG GAC TTC CAG AAG ACA ATG AAA GTC ACA GGA GTA ACT
ACT CAG GGA GTA AAA TCT CTG CTT ACC AGC ATG TAT GTG
AAG GAG TTC CTC ATC TCC AGC AGT CAA GAT GGC CAT CAG
TGG ACT CTC TTT TTT CAG AAT GGC AAA GTA AAG GTT TTT
CAG GGA AAT CAA GAC TCC TTC ACA CCT GTG GTG AAC TCT
CTA GAC CCA CCG TTA CTG ACT CGC TAC CTT CGA ATT CAC
CCC CAG AGT TGG GTG CAC CAG ATT GCC CTG AGG ATG GAG
GTT CTG GGC TGC GAG GCA CAG GAC CTC TAC; and
GCC ACC AGA AGA TAC TAC CTG GGT GCA GTG GAA CTG TCA
TGG GAC TAT ATG CAA AGT GAT CTC GGT GAG CTG CCT GTG
GAC GCA AGA TTT CCT CCT AGA GTG CCA AAA TCT TTT CCA
TTC AAC ACC TCA GTC GTG TAC AAA AAG ACT CTG TTT GTA
GAA TTC ACG GAT CAC CTT TTC AAC ATC GCT AAG CCA AGG
CCA CCC TGG ATG GGT CTG CTA GGT CCT ACC ATC CAG GCT
GAG GTT TAT GAT ACA GTG GTC ATT ACA CTT AAG AAC ATG
GCT TCC CAT CCT GTC AGT CTT CAT GCT GTT GGT GTA TCC

TAC TGG AAA GCT TCT GAG GGA GCT GAA TAT GAT GAT CAG
ACC AGT CAA AGG GAG AAA GAA GAT GAT AAA GTC TTC CCT
GGT GGA AGC CAT ACA TAT GTC TGG CAG GTC CTG AAA GAG
AAT GGT CCA ATG GCC TCT GAC CCA CTG TGC CTT ACC TAC
TCA TAT CTT TCT CAT GTG GAC CTG GTA AAA GAC TTG AAT
TCA GGC CTC ATT GGA GCC CTA CTA GTA TGT AGA GAA GGG
AGT CTG GCC AAG GAA AAG ACA CAC ACC TTG CAC AAA TTT
ATA CTA CTT TTT GCT GTA TTT GAT GAA GGG AAA AGT TGG
CAC TCA GAA ACA AAG AAC TCC TTG ATG CAG GAT AGG GAT
GCT GCA TCT GCT CGG GCC TGG CCT AAA ATG CAC ACA GTC
AAT GGT TAT GTA AAC AGG TCT CTG(CTA) CCA GGT CTG ATT
GGA TGC CAC AGG AAA TCA GTC TAT TGG CAT GTG ATT GGA
ATG GGC ACC ACT CCT GAA GTG CAC TCA ATA TTC CTC GAA
GGT CAC ACA TTT CTT GTG AGG AAC CAT CGC CAG GCG TCC
TTG GAA ATC TCG CCA ATA ACT TTC CTT ACT GCT CAA ACA
CTC TTG ATG GAC CTT GGA CAG TTT CTA CTG TTT TGT CAT
ATC TCT TCC CAC CAA CAT GAT GGC ATG GAA GCT TAT GTC
AAA GTA GAC AGC TGT CCA GAG GAA CCC CAA CTA CGA ATG
AAA AAT AAT GAA GAA GCG GAA GAC TAT GAT GAT GAT CTT
ACT GAT TCT GAA ATG GAT GTG GTC AGG TTT GAT GAT GAC
AAC TCT CCT TCC TTT ATC CAA ATT CGC TCA GTT GCC AAG
AAG CAT CCT AAA ACT TGG GTA CAT TAC ATT GCT GCT GAA
GAG GAG GAC TGG GAC TAT GCT CCC TTA GTC CTC GCC CCC
GAT GAC AGA AGT TAT AAA AGT CAA TAT TTG AAC AAT GGC
CCT CAG CGG ATT GGT AGG AAG TAC AAA AAA GTC CGA TTT
ATG GCA TAC ACA GAT GAA ACC TTT AAG ACT CGT GAA GCT
ATT CAG CAT GAA TCA GGA ATC TTG GGA CCT TTA CTT TAT
GGG GAA GTT GGA GAC ACA CTG TTG ATT ATA TTT AAG AAT
CAA GCA AGC AGA CCA TAT AAC ATC TAC CCT CAC GGA ATC
ACT GAT GTC CGT CCT TTG TAT TCA AGG AGA TTA CCA AAA
GGT GTA AAA CAT TTG AAG GAT TTT CCA ATT CTG CCA GGA
GAA ATA TTC AAA TAT AAA TGG ACA GTG ACT GTA GAA GAT
GGG CCA ACT AAA TCA GAT CCT CGG TGC CTG ACC CGC TAT
TAC TCT AGT TTC GTT AAT ATG GAG AGA GAT CTA GCT TCA
GGA CTC ATT GGC CCT CTC CTC ATC TGC TAC AAA GAA TCT
GTA GAT CAA AGA GGA AAC CAG ATA ATG TCA GAC AAG AGG
AAT GTC ATC CTG TTT TCT GTA TTT GAT GAG AAC CGA AGC

TGG TAC CTC ACA GAG AAT ATA CAA CGC TTT CTC CCC AAT
CCA GCT GGA GTG CAG CTT GAG GAT CCA GAG TTC CAA GCC
TCC AAC ATC ATG CAC AGC ATC AAT GGC TAT GTT TTT GAT
AGT TTG CAG TTG TCA GTT TGT TTG CAT GAG GTG GCA TAC
TGG TAC ATT CTA AGC ATT GGA GCA CAG ACT GAC TTC CTT
TCT GTC TTC TTC TCT GGA TAT ACC TTC AAA CAC AAA ATG
GTC TAT GAA GAC ACA CTC ACC CTA TTC CCA TTC TCA GGA
GAA ACT GTC TTC ATG TCG ATG GAA AAC CCA GGT CTA TGG
ATT CTG GGG TGC CAC AAC TCA GAC TTT CGG AAC AGA GGC
ATG ACC GCC TTA CTG AAG GTT TCT AGT TGT GAC AAG AAC
ACT GGT GAT TAT TAC GAG GAC AGT TAT GAA GAT ATT TCA
GCA TAC TTG CTG AGT AAA AAC AAT GCC ATT GAA CCA AGA
GAA ATA ACT CGT ACT ACT CTT CAG TCA GAT CAA GAG GAA
ATT GAC TAT GAT GAT ACC ATA TCA GTT GAA ATG AAG AAG
GAA GAT TTT GAC ATT TAT GAT GAG GAT GAA AAT CAG AGC
CCC CGC AGC TTT CAA AAG AAA ACA CGA CAC TAT TTT ATT
GCT GCA GTG GAG AGG CTC TGG GAT TAT GGG ATG AGT AGC
TCC CCA CAT GTT CTA AGA AAC AGG GCT CAG AGT GGC AGT
GTC CCT CAG TTC AAG AAA GTT GTT TTC CAG GAA TTT ACT
GAT GGC TCC TTT ACT CAG CCC TTA TAC CGT GGA GAA CTA
AAT GAA CAT TTG GGA CTC CTG GGG CCA TAT ATA AGA GCA
GAA GTT GAA GAT AAT ATC ATG GTA ACT TTC AGA AAT CAG
GCC TCT CGT CCC TAT TCC TTC TAT TCT AGC CTT ATT TCT
TAT GAG GAA GAT CAG AGG CAA GGA GCA GAA CCT AGA AAA
AAC TTT GTC AAG CCT AAT GAA ACC AAA ACT TAC TTT TGG
AAA GTG CAA CAT CAT ATG GCA CCC ACT AAA GAT GAG TTT
GAC TGC AAA GCC TGG GCT TAT TTC TCT GAT GTT GAC CTG
GAA AAA GAT GTG CAC TCA GGC CTG ATT GGA CCC CTT CTG
GTC TGC CAC ACT AAC ACA CTG AAC CCT GCT CAT GGG AGA
CAA GTG ACA GTA CAG GAA TTT GCT CTG TTT TTC(CTC) ACC
ATC TTT GAT GAG ACC AAA AGC TGG TAC TTC ACT GAA AAT
ATG GAA AGA AAC TGC AGG GCT CCC TGC AAT ATC CAG ATG
GAA GAT CCC ACT TTT AAA GAG AAT TAT CGC TTC CAT GCA
ATC AAT GGC TAC ATA ATG GAT ACA CTA CCT GGC TTA GTA
ATG GCT CAG GAT CAA AGG ATT CGA TGG TAT CTG CTC AGC
ATG GGC AGC AAT GAA AAC ATC CAT TCT ATT CAT TTC AGT
GGA CAT GTG TTC ACT GTA CGA AAA AAA GAG GAG TAT AAA

ATG GCA CTG TAC AAT CTC TAT CCA GGT GTT TTT GAG ACA
GTG GAA ATG TTA CCA TCC AAA GCT GGA ATT TGG CGG GTG
GAA TGC CTT ATT GGC GAG CAT CTA CAT GCT GGG ATG AGC
ACA CTT TTT CTG GTG TAC AGC AAT AAG TGT CAG ACT CCC
CTG GGA ATG GCT TCT GGA CAC ATT AGA GAT TTT CAG ATT
ACA GCT TCA GGA CAA TAT GGA CAG TGG GCC CCA AAG CTG
GCC AGA CTT CAT TAT TCC GGA TCA ATC AAT GCC TGG AGC
ACC AAG GAG CCC TTT TCT TGG ATC AAG GTG GAT CTG TTG
GCA CCA ATG ATT ATT CAC GGC ATC AAG ACC CAG GGT GCC
CGT CAG AAG TTC TCC AGC CTC TAC ATC TCT CAG TTT ATC
ATC ATG TAT AGT CTT GAT GGG AAG AAG TGG CAG ACT TAT
CGA GGA AAT TCC ACT GGA ACC TTA ATG GTC TTC TTT GGC
AAT GTG GAT TCA TCT GGG ATA AAA CAC AAT ATT TTT AAC
CCT CCA ATT ATT GCT CGA TAC ATC CGT TTG CAC CCA ACT
CAT TAT AGC ATT CGC AGC ACT CTT CGC ATG GAG TTG ATG
GGC TGT GAT TTA AAT AGT TGC AGC ATG CCA TTG GGA ATG
GAG AGT AAA GCA ATA TCA GAT GCA CAG ATT ACT GCT TCA
TCC TAC TTT ACC AAT ATG TTT GCC ACC TGG TCT CCT TCA
AAA GCT CGA CTT CAC CTC CAA GGG AGG AGT AAT GCC TGG
AGA CCT CAG GTG AAT AAT CCA AAA GAG TGG CTG CAA GTG
GAC TTC CAG AAG ACA ATG AAA GTC ACA GGA GTA ACT ACT
CAG GGA GTA AAA TCT CTG CTT ACC AGC ATG TAT GTG AAG
GAG TTC CTC ATC TCC AGC AGT CAA GAT GGC CAT CAG TGG
ACT CTC TTT TTT CAG AAT GGC AAA GTA AAG GTT TTT CAG
GGA AAT CAA GAC TCC TTC ACA CCT GTG GTG AAC TCT CTA
GAC CCA CCG TTA CTG ACT CGC TAC CTT CGA ATT CAC CCC
CAG AGT TGG GTG CAC CAG ATT GCC CTG AGG ATG GAG GTT
CTG GGC TGC GAG GCA CAG GAC CTC TAC.

4. The recombinant DNA molecule according to any one of claims 1-3, wherein the DNA sequence coding on expression for the modified factor VIII:C-like polypeptide is operatively linked to an expression control sequence.

5. The recombinant DNA molecule according to claim 4, wherein the expression control sequence is selected from the group consisting of the lac system, the trp system, the tac system, the trc system, major operator and promoter regions of phage λ, the control region of fd coat protein, the early and late promoters of SV40, promoters derived from polyoma, adenovirus and simian virus, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of yeast acid phosphatase, the promoters of the yeast α-mating factors, and other sequences known to control the expression of genes of prokaryotic or eukaryotic cells and their viruses, or combinations thereof.

6. A modified factor VIII:C-like polypeptide having a formula selected from the group consisting of:

met ala thr arg arg tyr tyr leu gly ala val glu leu
ser trp asp tyr met gln ser asp leu gly glu leu pro
val asp ala arg phe pro pro arg val pro lys ser phe
pro phe asn thr ser val val tyr lys lys thr leu phe
val glu phe thr asp his leu phe asn ile ala lys pro
arg pro pro trp met gly leu leu gly pro thr ile gln
ala glu val tyr asp thr val val ile thr leu lys asn
met ala ser his pro val ser leu his ala val gly val
ser tyr trp lys ala ser glu gly ala glu tyr asp asp
gln thr ser gln arg glu lys glu asp asp lys val phe
pro gly gly ser his thr tyr val trp gln val leu lys
glu asn gly pro met ala ser asp pro leu cys leu thr
tyr ser tyr leu ser his val asp leu val lys asp leu
asn ser gly leu ile gly ala leu leu val cys arg glu
gly ser leu ala lys glu lys thr gln thr leu his lys
phe ile leu leu phe ala val phe asp glu gly lys ser
trp his ser glu thr lys asn ser leu met gln asp arg
asp ala ala ser ala arg ala trp pro lys met his thr
val asn gly try val asn arg ser leu pro gly leu ile

gly cys his arg lys ser val tyr trp his val ile gly
met gly thr thr pro glu val his ser ile phe leu glu
gly his thr phe leu val arg asn his arg gln ala ser
leu glu ile ser pro ile thr phe leu thr ala gln thr
leu leu met asp leu gly gln phe leu leu phe cys his
ile ser ser his gln his asp gly met glu ala tyr val
lys val asp ser cys pro glu glu pro gln leu arg met
lys asn asn glu glu ala glu asp tyr asp asp asp leu
thr asp ser glu met asp val val arg phe asp asp asp
asn ser pro ser phe ile gln ile arg ser val ala lys
lys his pro lys thr trp val his tyr ile ala ala glu
glu glu asp trp asp tyr ala pro leu val leu ala pro
asp asp arg ser tyr lys ser gln tyr leu asn asn gly
pro gln arg ile gly arg lys tyr lys lys val arg phe
met ala tyr thr asp glu thr phe lys thr arg glu ala
ile gln his glu ser gly ile leu gly pro leu leu tyr
gly glu val gly asp thr leu leu ile ile phe lys asn
gln ala ser arg pro tyr asn ile tyr pro his gly ile
thr asp val arg pro leu tyr ser arg arg leu pro lys
gly val lys his leu lys asp phe pro ile leu pro gly
glu ile phe lys tyr lys trp thr val thr val glu asp
gly pro thr lys ser asp pro arg cys leu thr arg tyr
tyr ser ser phe val asn met glu arg asp leu ala ser
gly leu ile gly pro leu leu ile cys tyr lys glu ser
val asp gln arg gly asn gln ile met ser asp lys arg
asn val ile leu phe ser val phe asp glu asn arg ser
trp tyr leu thr glu asn ile gln arg phe leu pro asn
pro ala gly val gln leu glu asp pro glu phe gln ala
ser asn ile met his ser ile asn gly tyr val phe asp
ser leu gln leu ser val cys leu his glu val ala tyr
trp tyr ile leu ser ile gly ala gln thr asp phe leu
ser val phe phe ser gly tyr thr phe lys his lys met
val tyr glu asp thr leu thr leu phe pro phe ser gly
glu thr val phe met ser met glu asn pro gly leu trp
ile leu gly cys his asn ser asp phe arg asn arg gly
met thr ala leu leu lys val ser ser cys asp lys asn
thr gly asp tyr tyr glu asp ser tyr glu asp ile ser

ala tyr leu leu ser lys asn asn ala ile glu pro arg
ser phe ser gln asp pro leu ala trp asp asn his tyr
gly thr gln ile pro lys glu glu trp lys ser gln glu
lys ser pro glu lys thr ala phe lys lys lys asp thr
ile leu ser leu asn ala cys glu ser asn his ala ile
ala ala ile asn glu gly gln asn lys pro glu ile glu
val thr trp ala lys gln gly arg thr glu arg leu cys
ser gln asn pro pro val leu lys arg his gln arg glu
ile thr arg thr thr leu gln ser asp gln glu glu ile
asp tyr asp asp thr ile ser val glu met lys lys glu
asp phe asp ile tyr asp glu asp glu asn gln ser pro
arg ser phe gln lys lys thr arg his tyr phe ile ala
ala val glu arg leu trp asp tyr gly met ser ser ser
pro his val leu arg asn arg ala gln ser gly ser val
pro gln phe lys lys val val phe gln glu phe thr asp
gly ser phe thr gln pro leu tyr arg gly glu leu asn
glu his leu gly leu leu gly pro tyr ile arg ala glu
val glu asp asn ile met val thr phe arg asn gln ala
ser arg pro tyr ser phe tyr ser ser leu ile ser tyr
glu glu asp gln arg gln gly ala glu pro arg lys asn
phe val lys pro asn glu thr lys thr tyr phe trp lys
val gln his his met ala pro thr lys asp glu phe asp
cys lys ala trp ala tyr phe ser asp val asp leu glu
lys asp val his ser gly leu ile gly pro leu leu val
cys his thr asn thr leu asn pro ala his gly arg gln
val thr val gln glu phe ala leu phe phe(leu) thr
ile phe asp glu thr lys ser trp tyr phe thr glu asn
met glu arg asn cys arg ala pro cys asn ile gln met
glu asp pro thr phe lys glu asn tyr arg phe his ala
ile asn gly tyr ile met asp thr leu pro gly leu val
met ala gln asp gln arg ile arg trp tyr leu leu ser
met gly ser asn glu asn ile his ser ile his phe ser
gly his val phe thr val arg lys lys glu glu tyr lys
met ala leu tyr asn leu tyr pro gly val phe glu thr
val glu met leu pro ser lys ala gly ile trp arg val
glu cys leu ile gly glu his leu his ala gly met ser
thr leu phe leu val tyr ser asn lys cys gln thr pro

leu gly met ala ser gly his ile arg asp phe gln ile
thr ala ser gly gln tyr gly gln trp ala pro lys leu
ala arg leu his tyr ser gly ser ile asn ala trp ser
thr lys glu pro phe ser trp ile lys val asp leu leu
ala pro met ile ile his gly ile lys thr gln gly ala
arg gln lys phe ser ser leu tyr ile ser gln phe ile
ile met tyr ser leu asp gly lys lys trp gln thr tyr
arg gly asn ser thr gly thr leu met val phe phe gly
asn val asp ser ser gly ile lys his asn ile phe asn
pro pro ile ile ala arg tyr ile arg leu his pro thr
his tyr ser ile arg ser thr leu arg met glu leu met
gly cys asp leu asn ser cys ser met pro leu gly met
glu ser lys ala ile ser asp ala gln ile thr ala ser
ser tyr phe thr asn met phe ala thr trp ser pro ser
lys ala arg leu his leu gln gly arg ser asn ala trp
arg pro gln val asn asn pro lys glu trp leu gln val
asp phe gln lys thr met lys val thr gly val thr thr
gln gly val lys ser leu leu thr ser met tyr val lys
glu phe leu ile ser ser ser gln asp gly his gln trp
thr leu phe phe gln asn gly lys val lys val phe gln
gly asn gln asp ser phe thr pro val val asn ser leu
asp pro pro leu leu thr arg tyr leu arg ile his pro
gln ser trp val his gln ile ala leu arg met glu val
leu gly cys glu ala gln asp leu tyr;
ala thr arg arg tyr tyr leu gly ala val glu leu ser
trp asp tyr met gln ser asp leu gly glu leu pro val
asp ala arg phe pro pro arg val pro lys ser phe pro
phe asn thr ser val val tyr lys lys thr leu phe val
glu phe thr asp his leu phe asn ile ala lys pro arg
pro pro trp met gly leu leu gly pro thr ile gln ala
glu val tyr asp thr val val ile thr leu lys asn met
ala ser his pro val ser leu his ala val gly val ser
tyr trp lys ala ser glu gly ala glu tyr asp asp gln
thr ser gln arg glu lys glu asp asp lys val phe pro
gly gly ser his thr tyr val trp gln val leu lys glu
asn gly pro met ala ser asp pro leu cys leu thr tyr
ser tyr leu ser his val asp leu val lys asp leu asn

ser gly leu ile gly ala leu leu val cys arg glu gly
ser leu ala lys glu lys thr gln thr leu his lys phe
ile leu leu phe ala val phe asp glu gly lys ser trp
his ser glu thr lys asn ser leu met gln asp arg asp
ala ala ser ala arg ala trp pro lys met his thr val
asn gly try val asn arg ser leu pro gly leu ile gly
cys his arg lys ser val tyr trp his val ile gly met
gly thr thr pro glu val his ser ile phe leu glu gly
his thr phe leu val arg asn his arg gln ala ser leu
glu ile ser pro ile thr phe leu thr ala gln thr leu
leu met asp leu gly gln phe leu leu phe cys his ile
ser ser his gln his asp gly met glu ala tyr val lys
val asp ser cys pro glu glu pro gln leu arg met lys
asn asn glu glu ala glu asp tyr asp asp asp leu thr
asp ser glu met asp val val arg phe asp asp asp asn
ser pro ser phe ile gln ile arg ser val ala lys lys
his pro lys thr trp val his tyr ile ala ala glu glu
glu asp trp asp tyr ala pro leu val leu ala pro asp
asp arg ser tyr lys ser gln tyr leu asn asn gly pro
gln arg ile gly arg lys tyr lys lys val arg phe met
ala tyr thr asp glu thr phe lys thr arg glu ala ile
gln his glu ser gly ile leu gly pro leu leu tyr gly
glu val gly asp thr leu leu ile ile phe lys asn gln
ala ser arg pro tyr asn ile tyr pro his gly ile thr
asp val arg pro leu tyr ser arg arg leu pro lys gly
val lys his leu lys asp phe pro ile leu pro gly glu
ile phe lys tyr lys trp thr val thr val glu asp gly
pro thr lys ser asp pro arg cys leu thr arg tyr tyr
ser ser phe val asn met glu arg asp leu ala ser gly
leu ile gly pro leu leu ile cys tyr lys glu ser val
asp gln arg gly asn gln ile met ser asp lys arg asn
val ile leu phe ser val phe asp glu asn arg ser trp
tyr leu thr glu asn ile gln arg phe leu pro asn pro
ala gly val gln leu glu asp pro glu phe gln ala ser
asn ile met his ser ile asn gly tyr val phe asp ser
leu gln leu ser val cys leu his glu val ala tyr trp
tyr ile leu ser ile gly ala gln thr asp phe leu ser

val phe phe ser gly tyr thr phe lys his lys met val
tyr glu asp thr leu thr leu phe pro phe ser gly glu
thr val phe met ser met glu asn pro gly leu trp ile
leu gly cys his asn ser asp phe arg asn arg gly met
thr ala leu leu lys val ser ser cys asp lys asn thr
gly asp tyr tyr glu asp ser tyr glu asp ile ser ala
tyr leu leu ser lys asn asn ala ile glu pro arg ser
phe ser gln asp pro leu ala trp asp asn his tyr gly
thr gln ile pro lys glu glu trp lys ser gln glu lys
ser pro glu lys thr ala phe lys lys lys asp thr ile
leu ser leu asn ala cys glu ser asn his ala ile ala
ala ile asn glu gly gln asn lys pro glu ile glu val
thr trp ala lys gln gly arg thr glu arg leu cys ser
gln asn pro pro val leu lys arg his gln arg glu ile
thr arg thr thr leu gln ser asp gln glu glu ile asp
tyr asp asp thr ile ser val glu met lys lys glu asp
phe asp ile tyr asp glu asp glu asn gln ser pro arg
ser phe gln lys lys thr arg his tyr phe ile ala ala
val glu arg leu trp asp tyr gly met ser ser ser pro
his val leu arg asn arg ala gln ser gly ser val pro
gln phe lys lys val val phe gln glu phe thr asp gly
ser phe thr gln pro leu tyr arg gly glu leu asn glu
his leu gly leu leu gly pro tyr ile arg ala glu val
glu asp asn ile met val thr phe arg asn gln ala ser
arg pro tyr ser phe tyr ser ser leu ile ser tyr glu
glu asp gln arg gln gly ala glu pro arg lys asn phe
val lys pro asn glu thr lys thr tyr phe trp lys val
gln his his met ala pro thr lys asp glu phe asp cys
lys ala trp ala tyr phe ser asp val asp leu glu lys
asp val his ser gly leu ile gly pro leu leu val cys
his thr asn thr leu asn pro ala his gly arg gln val
thr val gln glu phe ala leu phe phe(leu) thr ile
phe asp glu thr lys ser trp tyr phe thr glu asn met
glu arg asn cys arg ala pro cys asn ile gln met glu
asp pro thr phe lys glu asn tyr arg phe his ala ile
asn gly tyr ile met asp thr leu pro gly leu val met
ala gln asp gln arg ile arg trp tyr leu leu ser met

gly ser asn glu asn ile his ser ile his phe ser gly
his val phe thr val arg lys lys glu glu tyr lys met
ala leu tyr asn leu tyr pro gly val phe glu thr val
glu met leu pro ser lys ala gly ile trp arg val glu
cys leu ile gly glu his leu his ala gly met ser thr
leu phe leu val tyr ser asn lys cys gln thr pro leu
gly met ala ser gly his ile arg asp phe gln ile thr
ala ser gly gln tyr gly gln trp ala pro lys leu ala
arg leu his tyr ser gly ser ile asn ala trp ser thr
lys glu pro phe ser trp ile lys val asp leu leu ala
pro met ile ile his gly ile lys thr gln gly ala arg
gln lys phe ser ser leu tyr ile ser gln phe ile ile
met tyr ser leu asp gly lys lys trp gln thr tyr arg
gly asn ser thr gly thr leu met val phe phe gly asn
val asp ser ser gly ile lys his asn ile phe asn pro
pro ile ile ala arg tyr ile arg leu his pro thr his
tyr ser ile arg ser thr leu arg met glu leu met gly
cys asp leu asn ser cys ser met pro leu gly met glu
ser lys ala ile ser asp ala gln ile thr ala ser ser
tyr phe thr asn met phe ala thr trp ser pro ser lys
ala arg leu his leu gln gly arg ser asn ala trp arg
pro gln val asn asn pro lys glu trp leu gln val asp
phe gln lys thr met lys val thr gly val thr thr gln
gly val lys ser leu leu thr ser met tyr val lys glu
phe leu ile ser ser ser gln asp gly his gln trp thr
leu phe phe gln asn gly lys val lys val phe gln gly
asn gln asp ser phe thr pro val val asn ser leu asp
pro pro leu leu thr arg tyr leu arg ile his pro gln
ser trp val his gln ile ala leu arg met glu val leu
gly cys glu ala gln asp leu tyr;
met ala thr arg arg tyr tyr leu gly ala val glu leu
ser trp asp tyr met gln ser asp leu gly glu leu pro
val asp ala arg phe pro pro arg val pro lys ser phe
pro phe asn thr ser val val tyr lys lys thr leu phe
val glu phe thr asp his leu phe asn ile ala lys pro
arg pro pro trp met gly leu leu gly pro thr ile gln
ala glu val tyr asp thr val val ile thr leu lys asn

met ala ser his pro val ser leu his ala val gly val
ser tyr trp lys ala ser glu gly ala glu tyr asp asp
gln thr ser gln arg glu lys glu asp asp lys val phe
pro gly gly ser his thr tyr val trp gln val leu lys
glu asn gly pro met ala ser asp pro leu cys leu thr
tyr ser tyr leu ser his val asp leu val lys asp leu
asn ser gly leu ile gly ala leu leu val cys arg glu
gly ser leu ala lys glu lys thr gln thr leu his lys
phe ile leu leu phe ala val phe asp glu gly lys ser
trp his ser glu thr lys asn ser leu met gln asp arg
asp ala ala ser ala arg ala trp pro lys met his thr
val asn gly tyr val asn arg ser leu pro gly leu ile
gly cys his arg lys ser val tyr trp his val ile gly
met gly thr thr pro glu val his ser ile phe leu glu
gly his thr phe leu val arg asn his arg gln ala ser
leu glu ile ser pro ile thr phe leu thr ala gln thr
leu leu met asp leu gly gln phe leu leu phe cys his
ile ser ser his gln his asp gly met glu ala tyr val
lys val asp ser cys pro glu glu pro gln leu arg met
lys asn asn glu glu ala glu asp tyr asp asp asp leu
thr asp ser glu met asp val val arg phe asp asp asp
asn ser pro ser phe ile gln ile arg ser val ala lys
lys his pro lys thr trp val his tyr ile ala ala glu
glu glu asp trp asp tyr ala pro leu val leu ala pro
asp asp arg ser tyr lys ser gln tyr leu asn asn gly
pro gln arg ile gly arg lys try lys lys val arg phe
met ala tyr thr asp glu thr phe lys thr arg glu ala
ile gln his glu ser gly ile leu gly pro leu leu tyr
gly glu val gly asp thr leu leu ile ile phe lys asn
gln ala ser arg pro tyr asn ile tyr pro his gly ile
thr asp val arg pro leu tyr ser arg arg leu pro lys
gly val lys his leu lys asp phe pro ile leu pro gly
glu ile phe lys tyr lys trp thr val thr val glu asp
gly pro thr lys ser asp pro arg cys leu thr arg tyr
tyr ser ser phe val asn met glu arg asp leu ala ser
gly leu ile gly pro leu leu ile cys tyr lys glu ser
val asp gln arg gly asn gln ile met ser asp lys arg

asn val ile leu phe ser val phe asp glu asn arg ser
trp tyr leu thr glu asn ile gln arg phe leu pro asn
pro ala gly val gln leu glu asp pro glu phe gln ala
ser asn ile met his ser ile asn gly tyr val phe asp
ser leu gln leu ser val cys leu his glu val ala tyr
trp tyr ile leu ser ile gly ala gln thr asp phe leu
ser val phe phe ser gly tyr thr phe lys his lys met
val tyr glu asp thr leu thr leu phe pro phe ser gly
glu thr val phe met ser met glu asn pro gly leu trp
ile leu gly cys his asn ser asp phe arg asn arg gly
met thr ala leu leu lys val ser ser cys asp lys asn
thr gly asp tyr tyr glu asp ser tyr glu asp ile ser
ala tyr leu leu ser lys asn asn ala ile glu pro arg
glu ile thr arg thr thr leu gln ser asp gln glu glu
ile asp tyr asp asp thr ile ser val glu met lys lys
glu asp phe asp ile tyr asp glu asp glu asn gln ser
pro arg ser phe gln lys lys thr arg his tyr phe ile
ala ala val glu arg leu trp asp tyr gly met ser ser
ser pro his val leu arg asn arg ala gln ser gly ser
val pro gln phe lys lys val val phe gln glu phe thr
asp gly ser phe thr gln pro leu tyr arg gly glu leu
asn glu his leu gly leu leu gly pro tyr ile arg ala
glu val glu asp asn ile met val thr phe arg asn gln
ala ser arg pro tyr ser phe tyr ser ser leu ile ser
tyr glu glu asp gln arg gln gly ala glu pro arg lys
asn phe val lys pro asn glu thr lys thr tyr phe trp
lys val gln his his met ala pro thr lys asp glu phe
asp cys lys ala trp ala tyr phe ser asp val asp leu
glu lys asp val his ser gly leu ile gly pro leu leu
val cys his thr asn thr leu asn pro ala his gly arg
gln val thr val gln glu phe ala leu phe phe(leu)
thr ile phe asp glu thr lys ser trp tyr phe thr glu
asn met glu arg asn cys arg ala pro cys asn ile gln
met glu asp pro thr phe lys glu asn tyr arg phe his
ala ile asn gly tyr ile met asp thr leu pro gly leu
val met ala gln asp gln arg ile arg trp tyr leu leu
ser met gly ser asn glu asn ile his ser ile his phe

ser gly his val phe thr val arg lys lys glu glu tyr
lys met ala leu tyr asn leu tyr pro gly val phe glu
thr val glu met leu pro ser lys ala gly ile trp arg
val glu cys leu ile gly glu his leu his ala gly met
ser thr leu phe leu val tyr ser asn lys cys gln thr
pro leu gly met ala ser gly his ile arg asp phe gln
ile thr ala ser gly gln tyr gly gln trp ala pro lys
leu ala arg leu his tyr ser gly ser ile asn ala trp
ser thr lys glu pro phe ser trp ile lys val asp leu
leu ala pro met ile ile his gly ile lys thr gln gly
ala arg gln lys phe ser ser leu tyr ile ser gln phe
ile ile met tyr ser leu asp gly lys lys trp gln thr
tyr arg gly asn ser thr gly thr leu met val phe phe
gly asn val asp ser ser gly ile lys his asn ile phe
asn pro pro ile ile ala arg tyr ile arg leu his pro
thr his tyr ser ile arg ser thr leu arg met glu leu
met gly cys asp leu asn ser cys ser met pro leu gly
met glu ser lys ala ile ser asp ala gln ile thr ala
ser ser tyr phe thr asn met phe ala thr trp ser pro
ser lys ala arg leu his leu gln gly arg ser asn ala
trp arg pro gln val asn asn pro lys glu trp leu gln
val asp phe gln lys thr met lys val thr gly val thr
thr gln gly val lys ser leu leu thr ser met tyr val
lys glu phe leu ile ser ser ser gln asp gly his gln
trp thr leu phe phe gln asn gly lys val lys val phe
gln gly asn gln asp ser phe thr pro val val asn ser
leu asp pro pro leu leu thr arg tyr leu arg ile his
pro gln ser trp val his gln ile ala leu arg met glu
val leu gly cys glu ala gln asp leu tyr; and
ala thr arg arg tyr tyr leu gly ala val glu leu ser
trp asp tyr met gln ser asp leu gly glu leu pro val
asp ala arg phe pro pro arg val pro lys ser phe pro
phe asn thr ser val val tyr lys lys thr leu phe val
glu phe thr asp his leu phe asn ile ala lys pro arg
pro pro trp met gly leu leu gly pro thr ile gln ala
glu val tyr asp thr val val ile thr leu lys asn met
ala ser his pro val ser leu his ala val gly val ser

tyr trp lys ala ser glu gly ala glu tyr asp asp gln
thr ser gln arg glu lys glu asp asp lys val phe pro
gly gly ser his thr tyr val trp gln val leu lys glu
asn gly pro met ala ser asp pro leu cys leu thr tyr
ser tyr leu ser his val asp leu val lys asp leu asn
ser gly leu ile gly ala leu leu val cys arg glu gly
ser leu ala lys glu lys thr gln thr leu his lys phe
ile leu leu phe ala val phe asp glu gly lys ser trp
his ser glu thr lys asn ser leu met gln asp arg asp
ala ala ser ala arg ala trp pro lys met his thr val
asn gly tyr val asn arg ser leu pro gly leu ile gly
cys his arg lys ser val tyr trp his val ile gly met
gly thr thr pro glu val his ser ile phe leu glu gly
his thr phe leu val arg asn his arg gln ala ser leu
glu ile ser pro ile thr phe leu thr ala gln thr leu
leu met asp leu gly gln phe leu leu phe cys his ile
ser ser his gln his asp gly met glu ala tyr val lys
val asp ser cys pro glu glu pro gln leu arg met lys
asn asn glu glu ala glu asp tyr asp asp asp leu thr
asp ser glu met asp val val arg phe asp asp asp asn
ser pro ser phe ile gln ile arg ser val ala lys lys
his pro lys thr trp val his tyr ile ala ala glu glu
glu asp trp asp tyr ala pro leu val leu ala pro asp
asp arg ser tyr lys ser gln tyr leu asn asn gly pro
gln arg ile gly arg lys try lys lys val arg phe met
ala tyr thr asp glu thr phe lys thr arg glu ala ile
gln his glu ser gly ile leu gly pro leu leu tyr gly
glu val gly asp thr leu leu ile ile phe lys asn gln
ala ser arg pro tyr asn ile tyr pro his gly ile thr
asp val arg pro leu tyr ser arg arg leu pro lys gly
val lys his leu lys asp phe pro ile leu pro gly glu
ile phe lys tyr lys trp thr val thr val glu asp gly
pro thr lys ser asp pro arg cys leu thr arg tyr tyr
ser ser phe val asn met glu arg asp leu ala ser gly
leu ile gly pro leu leu ile cys tyr lys glu ser val
asp gln arg gly asn gln ile met ser asp lys arg asn
val ile leu phe ser val phe asp glu asn arg ser trp

tyr leu thr glu asn ile gln arg phe leu pro asn pro
ala gly val gln leu glu asp pro glu phe gln ala ser
asn ile met his ser ile asn gly tyr val phe asp ser
leu gln leu ser val cys leu his glu val ala tyr trp
tyr ile leu ser ile gly ala gln thr asp phe leu ser
val phe phe ser gly tyr thr phe lys his lys met val
tyr glu asp thr leu thr leu phe pro phe ser gly glu
thr val phe met ser met glu asn pro gly leu trp ile
leu gly cys his asn ser asp phe arg asn arg gly met
thr ala leu leu lys val ser ser cys asp lys asn thr
gly asp tyr tyr glu asp ser tyr glu asp ile ser ala
tyr leu leu ser lys asn asn ala ile glu pro arg glu
ile thr arg thr thr leu gln ser asp gln glu glu ile
asp tyr asp asp thr ile ser val glu met lys lys glu
asp phe asp ile tyr asp glu asp glu asn gln ser pro
arg ser phe gln lys lys thr arg his tyr phe ile ala
ala val glu arg leu trp asp tyr gly met ser ser ser
pro his val leu arg asn arg ala gln ser gly ser val
pro gln phe lys lys val val phe gln glu phe thr asp
gly ser phe thr gln pro leu tyr arg gly glu leu asn
glu his leu gly leu leu gly pro tyr ile arg ala glu
val glu asp asn ile met val thr phe arg asn gln ala
ser arg pro tyr ser phe tyr ser ser leu ile ser tyr
glu glu asp gln arg gln gly ala glu pro arg lys asn
phe val lys pro asn glu thr lys thr tyr phe trp lys
val gln his his met ala pro thr lys asp glu phe asp
cys lys ala trp ala tyr phe ser asp val asp leu glu
lys asp val his ser gly leu ile gly pro leu leu val
cys his thr asn thr leu asn pro ala his gly arg gln
val thr val gln glu phe ala leu phe phe(leu) thr
ile phe asp glu thr lys ser trp tyr phe thr glu asn
met glu arg asn cys arg ala pro cys asn ile gln met
glu asp pro thr phe lys glu asn tyr arg phe his ala
ile asn gly tyr ile met asp thr leu pro gly leu val
met ala gln asp gln arg ile arg trp tyr leu leu ser
met gly ser asn glu asn ile his ser ile his phe ser
gly his val phe thr val arg lys lys glu glu tyr lys

met ala leu tyr asn leu tyr pro gly val phe glu thr
val glu met leu pro ser lys ala gly ile trp arg val
glu cys leu ile gly glu his leu his ala gly met ser
thr leu phe leu val tyr ser asn lys cys gln thr pro
leu gly met ala ser gly his ile arg asp phe gln ile
thr ala ser gly gln tyr gly gln trp ala pro lys leu
ala arg leu his tyr ser gly ser ile asn ala trp ser
thr lys glu pro phe ser trp ile lys val asp leu leu
ala pro met ile ile his gly ile lys thr gln gly ala
arg gln lys phe ser ser leu tyr ile ser gln phe ile
ile met tyr ser leu asp gly lys lys trp gln thr tyr
arg gly asn ser thr gly thr leu met val phe phe gly
asn val asp ser ser gly ile lys his asn ile phe asn
pro pro ile ile ala arg tyr ile arg leu his pro thr
his tyr ser ile arg ser thr leu arg met glu leu met
gly cys asp leu asn ser cys ser met pro leu gly met
glu ser lys ala ile ser asp ala gln ile thr ala ser
ser tyr phe thr asn met phe ala thr trp ser pro ser
lys ala arg leu his leu gln gly arg ser asn ala trp
arg pro gln val asn asn pro lys glu trp leu gln val
asp phe gln lys thr met lys val thr gly val thr thr
gln gly val lys ser leu leu thr ser met tyr val lys
glu phe leu ile ser ser ser gln asp gly his gln trp
thr leu phe phe gln asn gly lys val lys val phe gln
gly asn gln asp ser phe thr pro val val asn ser leu
asp pro pro leu leu thr arg tyr leu arg ile his pro
gln ser trp val his gln ile ala leu arg met glu val
leu gly cys glu ala gln asp leu tyr.

7. A modified factor VIII:C-like polypeptide, comprising the N-terminal heavy chain of mature factor VIII:C linked directly to the C-terminal light chain of mature factor VIII:C, said polypeptide being essentially free of other serum proteins.

8. A process for producing a polypeptide, comprising the step of proteolytically cleaving the

modified factor VIII:C-like polypeptide of claim 7 into the N-terminal heavy chain of mature factor VIII:C and the C-terminal light chain of mature factor VIII:C.

9. The process according to claim 8, further comprising the step of linking together by an alkaline metal bridge, the N-terminal heavy chain of mature factor VIII:C and the C-terminal heavy chain of mature factor VIII:C.

10. A process for producing a modified factor VIII:C-like polypeptide comprising the step of culturing a host transformed with a recombinant DNA molecule as defined in claims 1 through 5.

11. The process according to any of claims 8, 9 or 10, wherein the host is selected from BMT10, BSC1, BSC40, COS1, COS7, CHO cells and other animal and human cells in culture.

12. A pharmaceutical composition comprising a polypeptide, produced according to the process of any of claims 8, 9 or 10, in an amount effective as a coagulant and a pharmaceutically acceptable carrier.

13. A pharmaceutical composition comprising a modified factor VIII:C-like polypeptide as defined in any one of claims 6-9 in an amount effective as a coagulant and a pharmaceutically acceptable carrier.

14. A method for treating haemophilia comprising the step of treating a human with the pharmaceutical composition as defined in claims 12 and 13.

15. A recombinant DNA molecule according to claim 4, selected from the group consisting of the recombinant DNA molecules contained in transformed host E.coli HB101(RE), E.coli (HB101(RD) or E.coli HB101(RSD).

16. A modified factor VIII:C-like polypeptide produced by a host transformed with a recombinant DNA molecule selected from a group consisting of recombinant DNA molecules contained in transformed host E.coli HB101(RE), E.coli HB101(RD) and E.coli HB101(RSD).

17. A process for producing a modified factor VIII:C-like polypeptide comprising the step of culturing a host transformed with a recombinant DNA molecule selected from a group consisting of recombinant DNA molecules contained in transformed host E.coli HB101(RE), E.coli HB101(RD) and E.coli HB101(RSD).

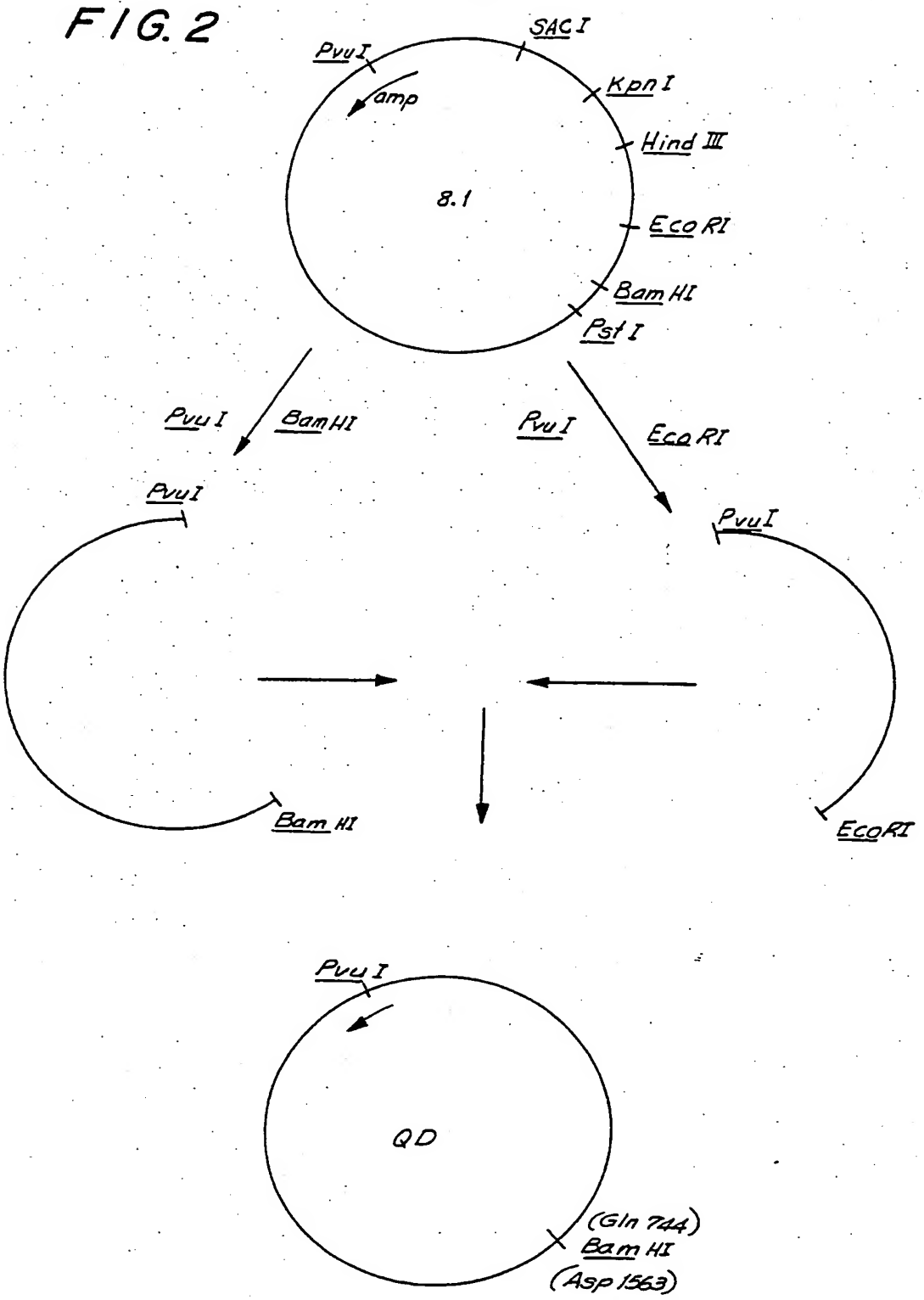
18. A pharmaceutical composition comprising a polypeptide produced according to the process of claim 17 in an amount effective as a coagulant and a pharmaceutically acceptable carrier.

19. A method for treating haemophilia comprising the step of treating a human with a pharmaceutical composition according to claim 18.

RESTRICTION ENDONUCLEASE MAP
OF THE FACTOR VIII:C cDNA
Ca+2
pUC 19 2874
1.7977
10.797783
7.74475
4.7573
2.73

FIG. 1

FIG. 2


PvuI
SAC I
Kpn I
amp
Hind III
8.1
Eco RI
Bam HI
Pst I
PvuI
Bam HI
PvuI
Eco RI
PvuI
Bam HI
PvuI
Eco RI
PvuI
QD
(Gln 744)
Bam HI
(Asp 1563)

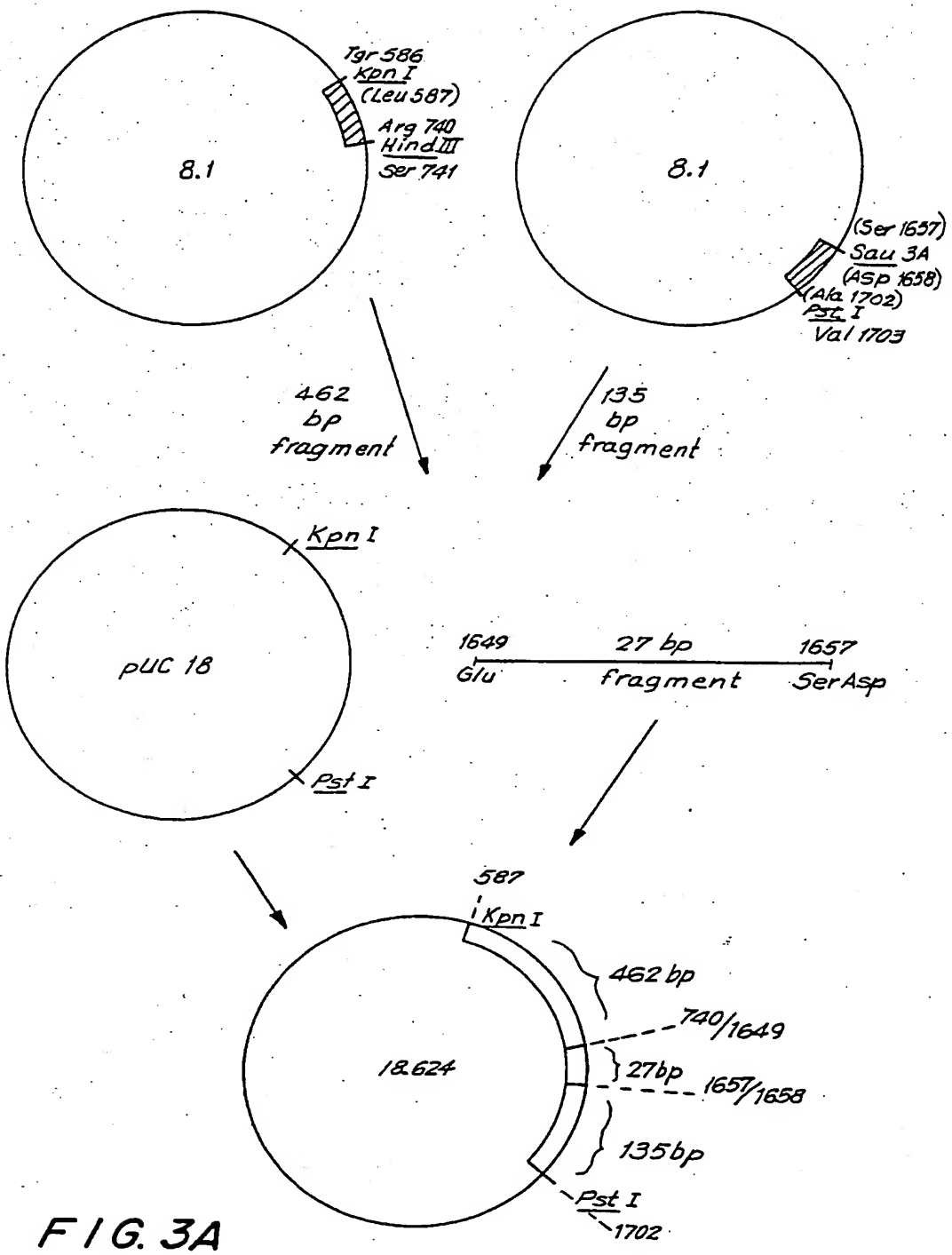

8.1
Tgr 586
Kpn I
(Leu 587)
Arg 740
Hind III
Ser 741
8.1
(Ser 1657)
Sau 3A
(Asp 1658)
(Ala 1702)
Pst I
Val 1703
462 bp fragment
135 bp fragment
pUC 18
Kpn I
Pst I
1649
27 bp
1657
Glu
fragment
SerAsp
587
Kpn I
462 bp
740/1649
18.624
27bp
1657/1658
135bp
Pst I
1702

FIG. 3A

FIG. 3B

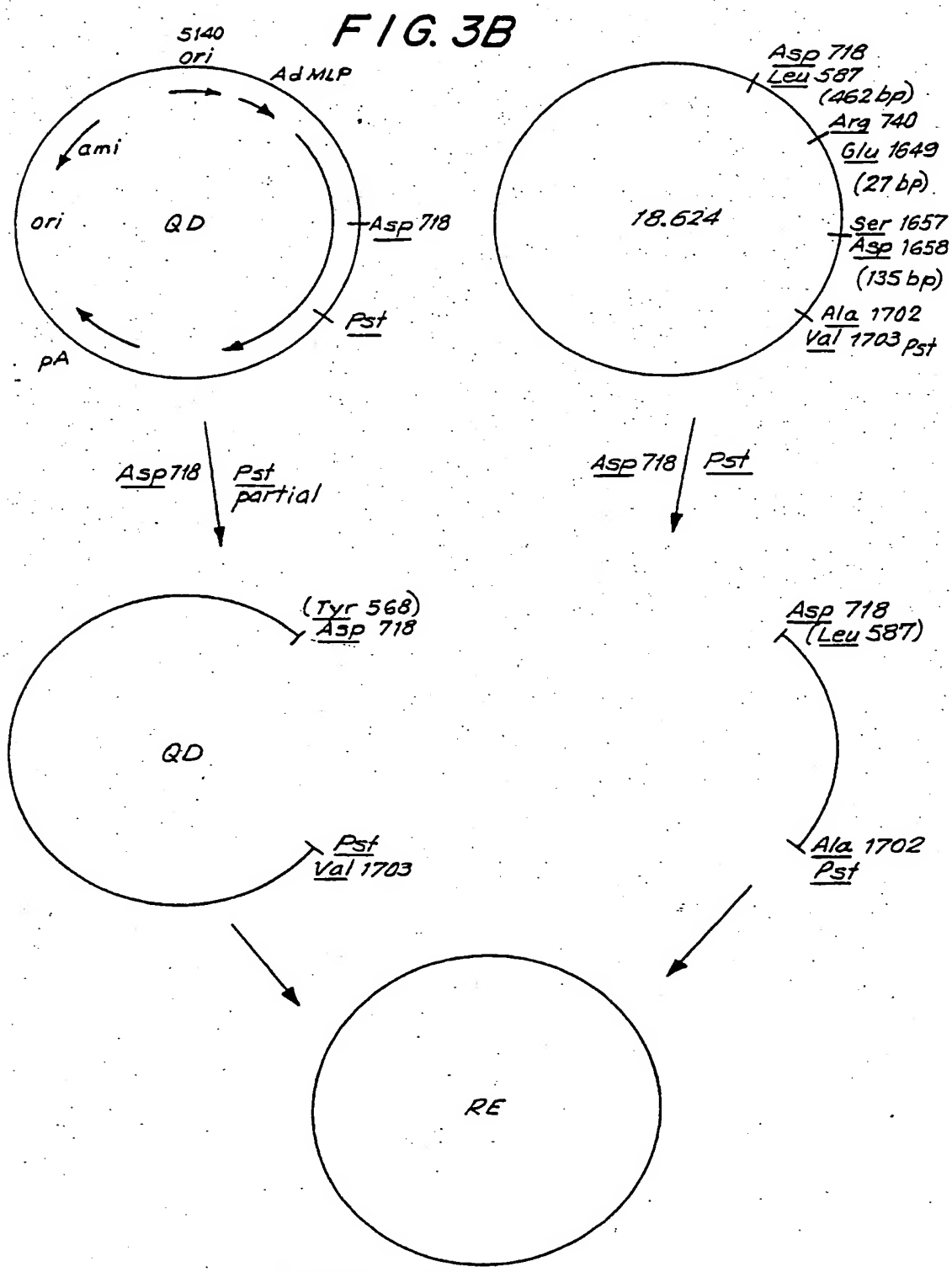
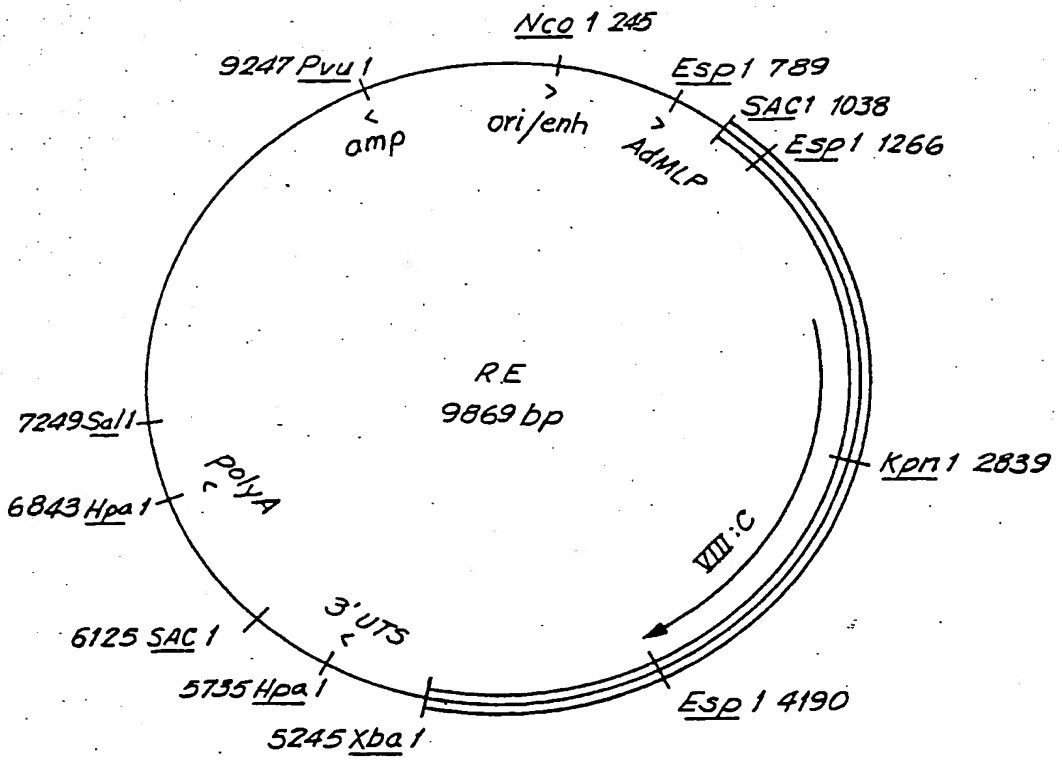

5140
ori
Ad MLP
ami
ori
QD
Asp 718
Pst
pA
Asp 718
Leu 587
(462 bp)
Arg 740
Glu 1649
(27 bp)
18.624
Ser 1657
Asp 1658
(135 bp)
Ala 1702
Val 1703 Pst
Asp 718
Pst
partial
Asp 718
Pst
(Tyr 568)
Asp 718
QD
Pst
Val 1703
Asp 718
(Leu 587)
Ala 1702
Pst
RE

FIG. 4


Nco 1 245
9247 Pvu 1
Esp 1 789
SAC1 1038
Esp 1 1266
amp
ori/enh
AdMLP
R.E
9869 bp
7249 Sal1
6843 Hpa 1
Poly A
Kpn 1 2839
VIII:C
3' UTS
6125 SAC 1
5735 Hpa 1
5245 Xba 1
Esp 1 4190

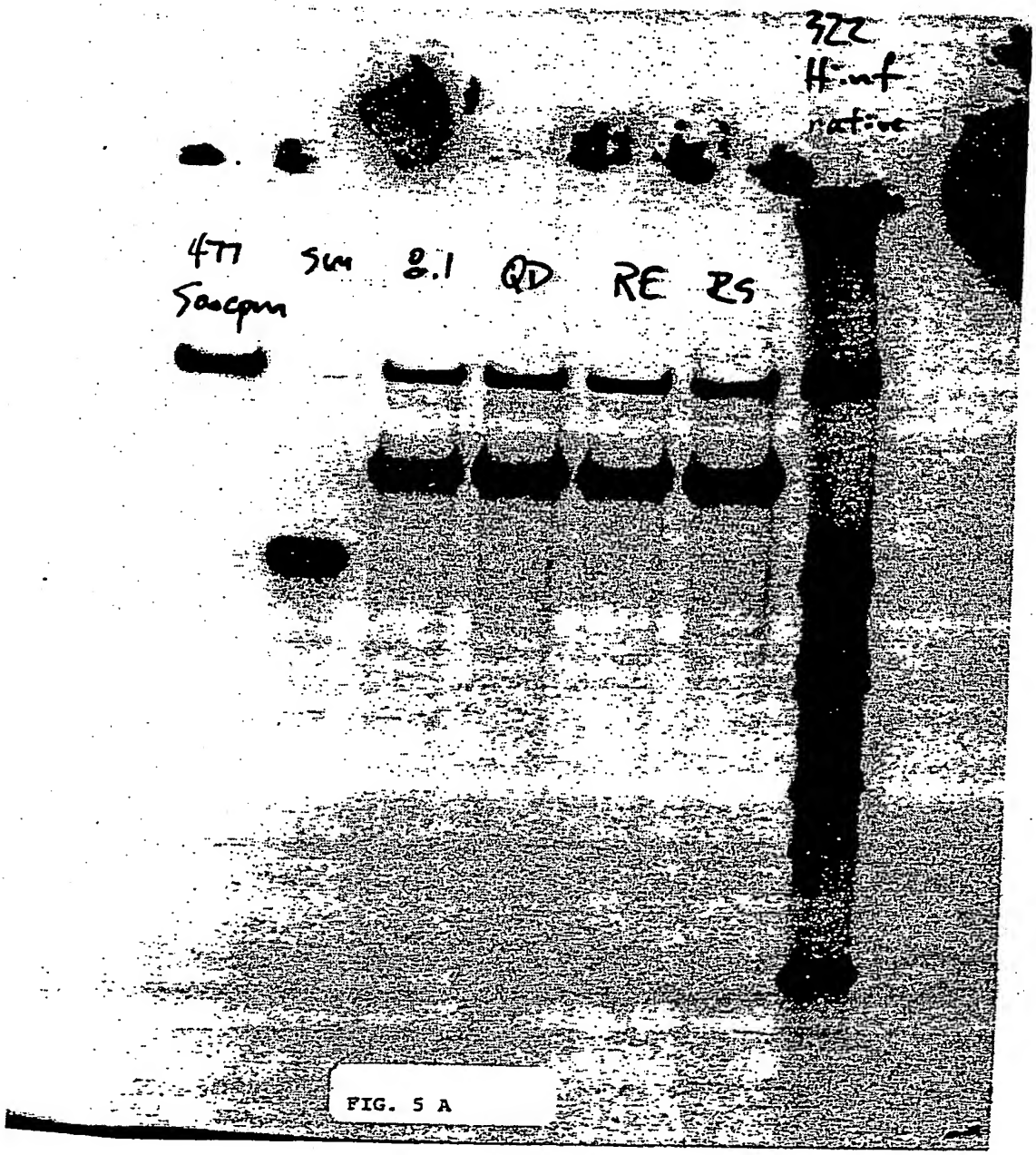


FIG. 5 A

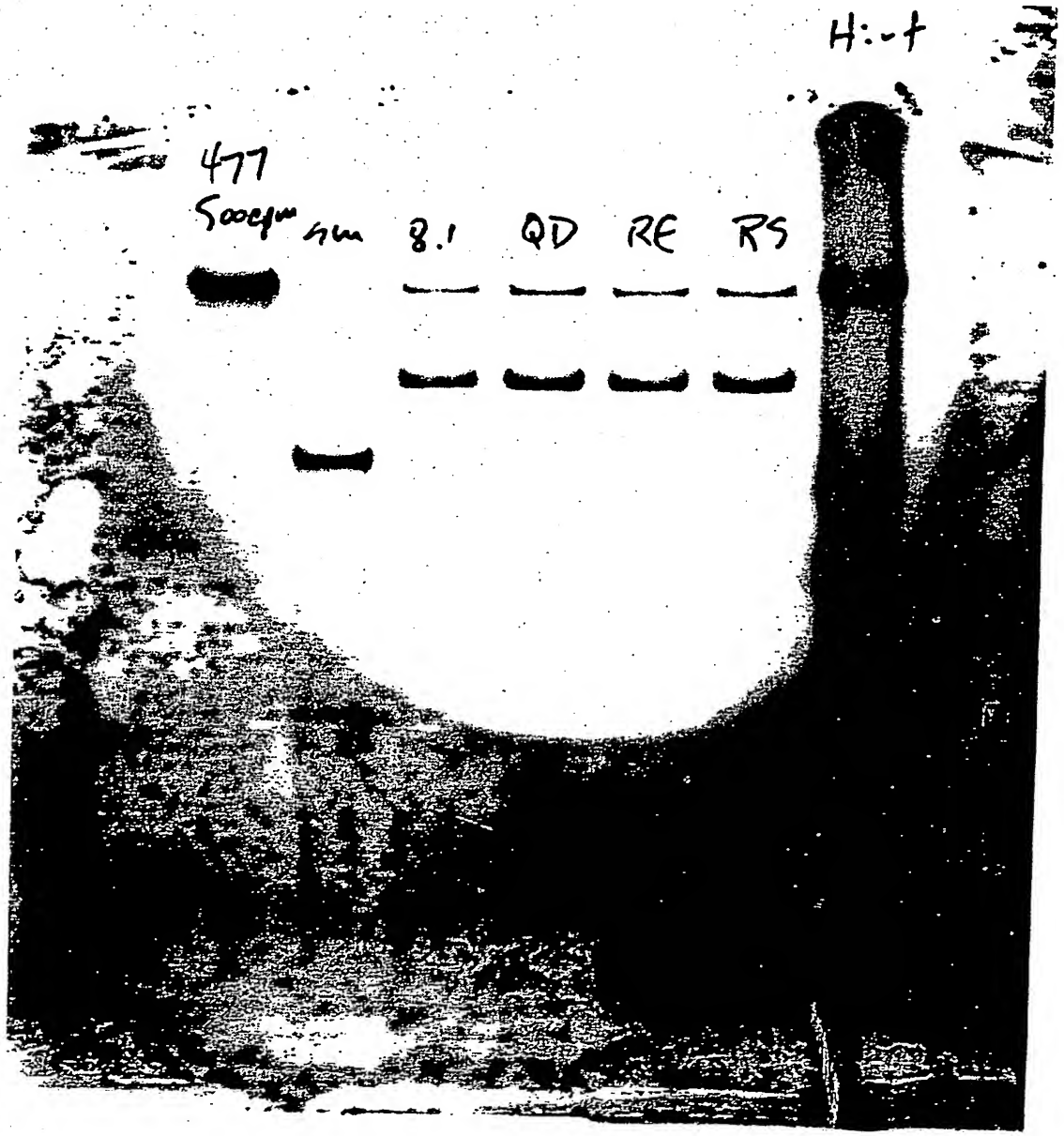


FIG. 5 B

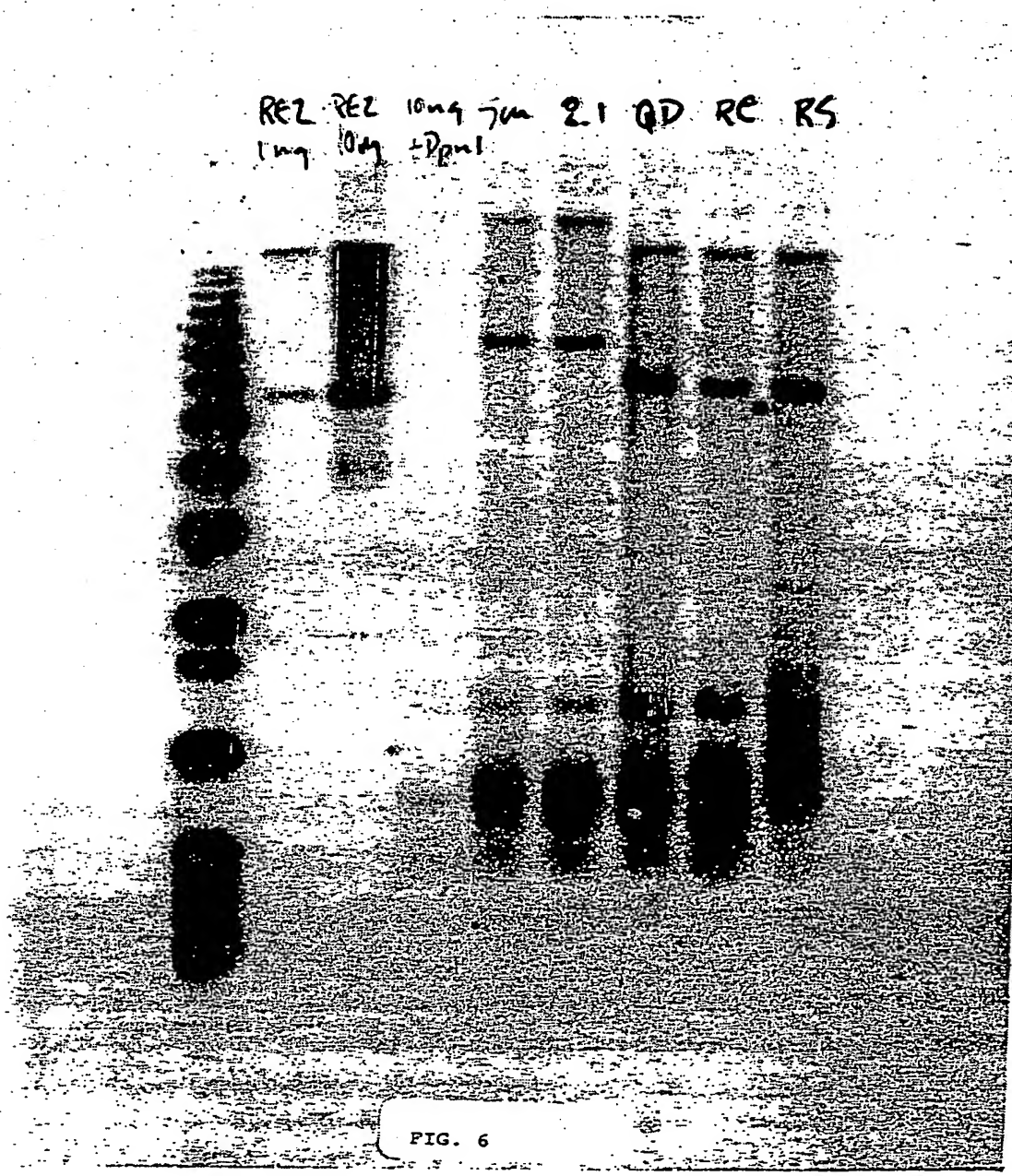


FIG. 6

```
  1                                                  10
 ala thr arg arg tyr tyr leu gly ala val glu leu ser
 GCC ACC AGA AGA TAC TAC CTG GGT GCA GTG GAA CTG TCA

                              20
 trp asp tyr met gln ser asp leu gly glu leu pro val
 TGG GAC TAT ATG CAA AGT GAT CTC GGT GAG CTG CCT GTG

                 30
 asp ala arg phe pro pro arg val pro lys ser phe pro
 GAC GCA AGA TTT CCT CCT AGA GTG CCA AAA TCT TTT CCA

 40                                                 50
 phe asn thr ser val val tyr lys lys thr leu phe val
 TTC AAC ACC TCA GTC GTG TAC AAA AAG ACT CTG TTT GTA

 ecoRI                          60
 glu phe thr asp his leu phe asn ile ala lys pro arg
 GAA TTC ACG GAT CAC CTT TTC AAC ATC GCT AAG CCA AGG

                    70
 pro pro trp met gly leu leu gly pro thr ile gln ala
 CCA CCC TGG ATG GGT CTG CTA GGT CCT ACC ATC CAG GCT

     80                                             90
 glu val tyr asp thr val val ile thr leu lys asn met
 GAG GTT TAT GAT ACA GTG GTC ATT ACA CTT AAG AAC ATG

                                   100
 ala ser his pro val ser leu his ala val gly val ser
 GCT TCC CAT CCT GTC AGT CTT CAT GCT GTT GGT GTA TCC
```

FIG. 7

```
      hindIII        110
tyr trp lys ala ser glu gly ala glu tyr asp asp gln
TAC TGG AAA GCT TCT GAG GGA GCT GAA TAT GAT GAT CAG

         120                                      130
thr ser gln arg glu lys glu asp asp lys val phe pro
ACC AGT CAA AGG GAG AAA GAA GAT GAT AAA GTC TTC CCT

                                    140
gly gly ser his thr tyr val trp gln val leu lys glu
GGT GGA AGC CAT ACA TAT GTC TGG CAG GTC CTG AAA GAG

                          150
asn gly pro met ala ser asp pro leu cys leu thr tyr
AAT GGT CCA ATG GCC TCT GAC CCA CTG TGC CTT ACC TAC

             160                                ecoRI
ser tyr leu ser his val asp leu val lys asp leu asn
TCA TAT CTT TCT CAT GTG GAC CTG GTA AAA GAC TTG AAT

170                                     180
ser gly leu ile gly ala leu leu val cys arg glu gly
TCA GGC CTC ATT GGA GCC CTA CTA GTA TGT AGA GAA GGG

                               190
ser leu ala lys glu lys thr gln thr leu his lys phe
AGT CTG GCC AAG GAA AAG ACA CAC ACC TTG CAC AAA TTT

                  200
ile leu leu phe ala val phe asp glu gly lys ser trp
ATA CTA CTT TTT GCT GTA TTT GAT GAA GGG AAA AGT TGG

    210                                       220
his ser glu thr lys asn ser leu met gln asp arg asp
CAC TCA GAA ACA AAG AAC TCC TTG ATG CAG GAT AGG GAT
```

FIG. 7(cont'd)

```
                                        230
ala ala ser ala arg ala trp pro lys met his thr val
GCT GCA TCT GCT CGG GCC TGG CCT AAA ATG CAC ACA GTC

                    240
asn gly try val asn arg ser leu pro gly leu ile gly
AAT GGT TAT GTA AAC AGG TCT CTG CCA GGT CTG ATT GGA

        250                                         260
cys his arg lys ser val tyr trp his val ile gly met
TGC CAC AGG AAA TCA GTC TAT TGG CAT GTG ATT GGA ATG

                                        270
gly thr thr pro glu val his ser ile phe leu glu gly
GGC ACC ACT CCT GAA GTG CAC TCA ATA TTC CTC GAA GGT

                        280
his thr phe leu val arg asn his arg gln ala ser leu
CAC ACA TTT CTT GTG AGG AAC CAT CGC CAG GCG TCC TTG

            290
glu ile ser pro ile thr phe leu thr ala gln thr leu
GAA ATC TCG CCA ATA ACT TTC CTT ACT GCT CAA ACA CTC

300                                         310
leu met asp leu gly gln phe leu leu phe cys his ile
TTG ATG GAC CTT GGA CAG TTT CTA CTG TTT TGT CAT ATC

                            320 hindIII
ser ser his gln his asp gly met glu ala tyr val lys
TCT TCC CAC CAA CAT GAT GGC ATG GAA GCT TAT GTC AAA

                    330
val asp ser cys pro glu glu pro gln leu arg met lys
GTA GAC AGC TGT CCA GAG GAA CCC CAA CTA CGA ATG AAA
```

FIG. 7(cont'd)

```
    340                                                 350
asn asn glu glu ala glu asp tyr asp asp asp leu thr
AAT AAT GAA GAA GCG GAA GAC TAT GAT GAT GAT CTT ACT

                                    360
asp ser glu met asp val val arg phe asp asp asp asn
GAT TCT GAA ATG GAT GTG GTC AGG TTT GAT GAT GAC AAC

                        370
ser pro ser phe ile gln ile arg ser val ala lys lys
TCT CCT TCC TTT ATC CAA ATT CGC TCA GTT GCC AAG AAG

        380                                         390
his pro lys thr trp val his tyr ile ala ala glu glu
CAT CCT AAA ACT TGG GTA CAT TAC ATT GCT GCT GAA GAG

                                        400
glu asp trp asp tyr ala pro leu val leu ala pro asp
GAG GAC TGG GAC TAT GCT CCC TTA GTC CTC GCC CCC GAT

                            410
asp arg ser tyr lys ser gln tyr leu asn asn gly pro
GAC AGA AGT TAT AAA AGT CAA TAT TTG AAC AAT GGC CCT

            420
gln arg ile gly arg lys tyr lys lys val arg phe met
CAG CGG ATT GGT AGG AAG TAC AAA AAA GTC CGA TTT ATG

430                                             440
ala tyr thr asp glu thr phe lys thr arg glu ala ile
GCA TAC ACA GAT GAA ACC TTT AAG ACT CGT GAA GCT ATT

                                    450
gln his glu ser gly ile leu gly pro leu leu tyr gly
CAG CAT GAA TCA GGA ATC TTG GGA CCT TTA CTT TAT GGG
```

FIG. 7(cont'd)

```
                460
glu val gly asp thr leu leu ile ile phe lys asn gln
GAA GTT GGA GAC ACA CTG TTG ATT ATA TTT AAG AAT CAA

   470                                          480
ala ser arg pro tyr asn ile tyr pro his gly ile thr
GCA AGC AGA CCA TAT AAC ATC TAC CCT CAC GGA ATC ACT

                                490
asp val arg pro leu tyr ser arg arg leu pro lys gly
GAT GTC CGT CCT TTG TAT TCA AGG AGA TTA CCA AAA GGT

                    500
val lys his leu lys asp phe pro ile leu pro gly glu
GTA AAA CAT TTG AAG GAT TTT CCA ATT CTG CCA GGA GAA

      510                                       520
ile phe lys tyr lys trp thr val thr val glu asp gly
ATA TTC AAA TAT AAA TGG ACA GTG ACT GTA GAA GAT GGG

                                    530
pro thr lys ser asp pro arg cys leu thr arg tyr tyr
CCA ACT AAA TCA GAT CCT CGG TGC CTG ACC CGC TAT TAC

                          540
ser ser phe val asn met glu arg asp leu ala ser gly
TCT AGT TTC GTT AAT ATG GAG AGA GAT CTA GCT TCA GGA

           550
leu ile gly pro leu leu ile cys tyr lys glu ser val
CTC ATT GGC CCT CTC CTC ATC TGC TAC AAA GAA TCT GTA

560                                         570
asp gln arg gly asn gln ile met ser asp lys arg asn
GAT CAA AGA GGA AAC CAG ATA ATG TCA GAC AAG AGG AAT
```

FIG. 7(cont'd)

SUBSTITUTE SHEET

```
                                        580                     k
val ile leu phe ser val phe asp glu asn arg ser trp
GTC ATC CTG TTT TCT GTA TTT GAT GAG AAC CGA AGC TGG

pnI             590
tyr leu thr glu asn ile gln arg phe leu pro asn pro
TAC CTC ACA GAG AAT ATA CAA CGC TTT CTC CCC AAT CCA

    600                 bamHI                   610
ala gly val gln leu glu asp pro glu phe gln ala ser
GCT GGA GTG CAG CTT GAG GAT CCA GAG TTC CAA GCC TCC

                                    620
asn ile met his ser ile asn gly tyr val phe asp ser
AAC ATC ATG CAC AGC ATC AAT GGC TAT GTT TTT GAT AGT

                    630
leu gln leu ser val cys leu his glu val ala tyr trp
TTG CAG TTG TCA GTT TGT TTG CAT GAG GTG GCA TAC TGG

        640                                         650
tyr ile leu ser ile gly ala gln thr asp phe leu ser
TAC ATT CTA AGC ATT GGA GCA CAG ACT GAC TTC CTT TCT

                                        660
val phe phe ser gly tyr thr phe lys his lys met val
GTC TTC TTC TCT GGA TAT ACC TTC AAA CAC AAA ATG GTC

                            670
tyr glu asp thr leu thr leu phe pro phe ser gly glu
TAT GAA GAC ACA CTC ACC CTA TTC CCA TTC TCA GGA GAA

            680
thr val phe met ser met glu asn pro gly leu trp ile
ACT GTC TTC ATG TCG ATG GAA AAC CCA GGT CTA TGG ATT
```

FIG. 7(cont'd)

690 700
leu gly cys his asn ser asp phe arg asn arg gly met
CTG GGG TGC CAC AAC TCA GAC TTT CGG AAC AGA GGC ATG

710
thr ala leu leu lys val ser ser cys asp lys asn thr
ACC GCC TTA CTG AAG GTT TCT AGT TGT GAC AAG AAC ACT

720
gly asp tyr tyr glu asp ser tyr glu asp ile ser ala
GGT GAT TAT TAC GAG GAC AGT TAT GAA GAT ATT TCA GCA

730 hindI
tyr leu leu ser lys asn asn ala ile glu pro arg ser
TAC TTG CTG AGT AAA AAC AAT GCC ATT GAA CCA AGA AGC

II ecoRI 750
phe ser gln asn ser arg his pro ser thr arg gln lys
TTC TCC CAG AAT TCA AGA CAC CCT AGC ACT AGG CAA AAG

760
gln phe asn ala thr thr ile pro glu asn asp ile glu
CAA TTT AAT GCC ACC ACA ATT CCA GAA AAT GAC ATA GAG

770 780
lys thr asp pro trp phe ala his arg thr pro met pro
AAG ACT GAC CCT TGG TTT GCA CAC AGA ACA CCT ATG CCT

790
lys ile gln asn val ser ser ser asp leu leu met leu
AAA ATA CAA AAT GTC TCC TCT AGT GAT TTG TTG ATG CTC

800
leu arg gln ser pro thr pro his gly leu ser leu ser
TTG CGA CAG AGT CCT ACT CCA CAT GGG CTA TCC TTA TCT

FIG. 7(cont'd)

```
                 810
asp leu gln glu ala lys tyr glu thr phe ser asp asp
GAT CTC CAA GAA GCC AAA TAT GAG ACT TTT TCT GAT GAT

820                                         830
pro ser pro gly ala ile asp ser asn asn ser leu ser
CCA TCA CCT GGA GCA ATA GAC AGT AAT AAC AGC CTG TCT

                                840
glu met thr his phe arg pro gln leu his his ser gly
GAA ATG ACA CAC TTC AGG CCA CAG CTC CAT CAC AGT GGG

                    850
asp met val phe thr pro glu ser gly leu gln leu arg
GAC ATG GTA TTT ACC CCT GAG TCA GGC CTC CAA TTA AGA

   860                                          870
leu asn glu lys leu gly thr thr ala ala thr glu leu
TTA AAT GAG AAA CTG GGG ACA ACT GCA GCA ACA GAG TTG

                                    880
lys lys leu asp phe lys val ser ser thr ser asn asn
AAG AAA CTT GAT TTC AAA GTT TCT AGT ACA TCA AAT AAT

                        890
leu ile ser thr ile pro ser asp asn leu ala ala gly
CTG ATT TCA ACA ATT CCA TCA GAC AAT TTG GCA GCA GGT

        900                                         910
thr asp asn thr ser ser leu gly pro pro ser met pro
ACT GAT AAT ACA AGT TCC TTA GGA CCC CCA AGT ATG CCA

                                        920
val his tyr asp ser gln leu asp thr thr leu phe gly
GTT CAT TAT GAT AGT CAA TTA GAT ACC ACT CTA TTT GGC
```

FIG. 7(cont'd)

```
                                930
 lys lys ser ser pro leu thr glu ser gly gly pro leu
 AAA AAG TCA TCT CCC CTT ACT GAG TCT GGT GGA CCT CTG

              940
 ser leu ser glu glu asn asn asp ser lys leu leu glu
 AGC TTG AGT GAA GAA AAT AAT GAT TCA AAG TTG TTA GAA

 950                                        960
 ser gly leu met asn ser gln glu ser ser trp gly lys
 TCA GGT TTA ATG AAT AGC CAA GAA AGT TCA TGG GGA AAA

                                970
 asn val ser ser thr glu ser gly arg leu phe lys gly
 AAT GTA TCG TCA ACA GAG AGT GGT AGG TTA TTT AAA GGG

      sacI       980
 lys arg ala his gly pro ala leu leu thr lys asp asn
 AAA AGA GCT CAT GGA CCT GCT TTG TTG ACT AAA GAT AAT

    990                                          1000
 ala leu phe lys val ser ile ser leu leu lys thr asn
 GCC TTA TTC AAA GTT AGC ATC TCT TTG TTA AAG ACA AAC

                                   1010
 lys thr ser asn asn ser ala thr asn arg lys thr his
 AAA ACT TCC AAT AAT TCA GCA ACT AAT AGA AAG ACT CAC

                         1020
 ile asp gly pro ser leu leu ile glu asn ser pro ser
 ATT GAT GGC CCA TCA TTA TTA ATT GAG AAT AGT CCA TCA

         1030                                     1040
 val trp gln asn ile leu glu ser asp thr glu phe lys
 GTC TGG CAA AAT ATA TTA GAA AGT GAC ACT GAG TTT AAA
```

FIG. 7(cont'd)

SUBSTITUTE SHEET

1050
lys val thr pro leu ile his asp arg met leu met asp
AAA GTG ACA CCT TTG ATT CAT GAC AGA ATG CTT ATG GAC

1060
lys asn ala thr ala leu arg leu asn his met ser asn
AAA AAT GCT ACA GCT TTG AGG CTA AAT CAT ATG TCA AAT

1070
lys thr thr ser ser lys asn met glu met val gln gln
AAA ACT ACT TCA TCA AAA AAC ATG GAA ATG GTC CAA CAG

1080 1090
lys lys glu gly pro ile pro pro asp ala gln asn pro
AAA AAA GAG GGC CCC ATT CCA CCA GAT GCA CAA AAT CCA

1100
asp met ser phe phe lys met leu phe leu pro glu ser
GAT ATG TCG TTC TTT AAG ATG CTA TTC TTG CCA GAA TCA

1110
ala arg trp ile gln arg thr his gly lys asn ser leu
GCA AGG TGG ATA CAA AGG ACT CAT GGA AAG AAC TCT CTG

1120 1130
asn ser gly gln gly pro ser pro lys gln leu val ser
AAC TCT GGG CAA GGC CCC AGT CCA AAG CAA TTA GTA TCC

1140
leu gly pro glu lys ser val glu gly gln asn phe leu
TTA GGA CCA GAA AAA TCT GTG GAA GGT CAG AAT TTC TTG

1150
ser glu lys asn lys val val val gly lys gly glu phe
TCT GAG AAA AAC AAA GTG GTA GTA GGA AAG GGT GAA TTT

FIG. 7(cont'd)

```
        1160                                                1170
thr lys asp val gly leu lys glu met val phe pro ser
ACA AAG GAC GTA GGA CTC AAA GAG ATG GTT TTT CCA AGC

                                            1180
ser arg asn leu phe leu thr asn leu asp asn leu his
AGC AGA AAC CTA TTT CTT ACT AAC TTG GAT AAT TTA CAT

                        1190
glu asn asn thr his asn gln glu lys lys ile gln glu
GAA AAT AAT ACA CAC AAT CAA GAA AAA AAA ATT CAG GAA

            1200
glu ile glu lys lys glu thr leu ile gln glu asn val
GAA ATA GAA AAG AAG GAA ACA TTA ATC CAA GAG AAT GTA

1210                                        1220
val leu pro gln ile his thr val thr gly thr lys asn
GTT TTG CCT CAG ATA CAT ACA GTG ACT GGC ACT AAG AAT

                                1230
phe met lys asn leu phe leu leu ser thr arg gln asn
TTC ATG AAG AAC CTT TTC TTA CTG AGC ACT AGG CAA AAT

                1240                            scaI
val glu gly ser tyr asp gly ala tyr ala pro val leu
GTA GAA GGT TCA TAT GAC GGG GCA TAT GCT CCA GTA CTT

    1250                                            1260
gln asp phe arg ser leu asn asp ser thr asn arg thr
CAA GAT TTT AGG TCA TTA AAT GAT TCA ACA AAT AGA ACA

                                    1270
lys lys his thr ala his phe ser lys lys gly glu glu
AAG AAA CAC ACA GCT CAT TTC TCA AAA AAA GGG GAG GAA
```

FIG. 7(cont'd)

```
                                1280
glu asn leu glu gly leu gly asn gln thr lys gln ile
GAA AAC TTG GAA GGC TTG GGA AAT CAA ACC AAG CAA ATT

        1290    sphI                            1300
val glu lys tyr ala cys thr thr arg ile ser pro asn
GTA GAG AAA TAT GCA TGC ACC ACA AGG ATA TCT CCT AAT

                                        1310
thr ser gln gln asn phe val thr gln arg ser lys arg
ACA AGC CAG CAG AAT TTT GTC ACG CAA CGT AGT AAG AGA

                            1320
ala leu lys gln phe arg leu pro leu glu glu thr glu
GCT TTG AAA CAA TTC AGA CTC CCA CTA GAA GAA ACA GAA

                1330
leu glu lys arg ile ile val asp asp thr ser thr gln
CTT GAA AAA AGG ATA ATT GTG GAT GAC ACC TCA ACC CAG

1340                                        1350
trp ser lys asn met lys his leu thr pro ser thr leu
TGG TCC AAA AAC ATG AAA CAT TTG ACC CCG AGC ACC CTC

                                1360
thr gln ile asp tyr asn glu lys glu lys gly ala ile
ACA CAG ATA GAC TAC AAT GAG AAG GAG AAA GGG GCC ATT

                    1370
thr gln ser pro leu ser asp cys leu thr arg ser his
ACT CAG TCT CCC TTA TCA GAT TCG CTT ACG AGG AGT CAT

    1380                                        1390
ser ile pro gln ala asn arg ser pro leu pro ile ala
AGC ATC CCT CAA GCA AAT AGA TCT CCA TTA CCC ATT GCA
```

FIG. 7(cont'd)

1400

lys val ser ser phe pro ser ile arg pro ile tyr leu
AAG GTA TCA TCA TTT CCA TCT ATT AGA CCT ATA TAT CTG

1410

thr arg val leu phe gln asp asn ser ser his leu pro
ACC AGG GTC CTA TTC CAA GAC AAC TCT TCT CAT CTT CCA

1420 1430

ala ala ser tyr arg lys lys asp ser gly val gln glu
GCA GCA TCT TAT AGA AAG AAA GAT TCT GGG GTC CAA GAA

1440

ser ser his phe leu gln gly ala lys lys asn asn leu
AGC AGT CAT TTC TTA CAA GGA GCC AAA AAA AAT AAC CTT

1450

ser leu ala ile leu thr leu glu met thr gly asp gln
TCT TTA GCC ATT CTA ACC TTG GAG ATG ACT GGT GAT CAA

1460

arg glu val gly ser leu gly thr ser ala thr asn ser
AGA GAG GTT GGC TCC CTG GGG ACA AGT GCC ACA AAT TCA

1470 1480

val thr tyr lys lys val glu asn thr val leu pro lys
GTC ACA TAC AAG AAA GTT GAG AAC ACT GTT CTC CCG AAA

1490

pro asp leu pro lys thr ser gly lys val glu leu leu
CCA GAC TTG CCC AAA ACA TCT GGC AAA GTT GAA TTG CTT

1500

pro lys val his ile tyr gln lys asp leu phe pro thr
CCA AAA GTT CAC ATT TAT CAG AAG GAC CTA TTC CCT ACG

FIG. 7(cont'd)

```
     1510                                                1520
 glu thr ser asn gly ser pro gly his leu asp leu val
 GAA ACT AGC AAT GGG TCT CCT GGC CAT CTG GAT CTC GTG

                                    1530
 glu gly ser leu leu gln gly thr glu gly ala ile lys
 GAA GGG AGC CTT CTT CAG GGA ACA GAG GGA GCG ATT AAG

                        1540
 trp asn glu ala asn arg pro gly lys val pro phe leu
 TGG AAT GAA GCA AAC AGA CCT GGA AAA GTT CCC TTT CTG

        1550                                          1560
 arg val ala thr glu ser ser ala lys thr pro ser lys
 AGA GTA GCA ACA GAA AGC TCT GCA AAG ACT CCC TCC AAG

      bamHI                             1570
 leu leu asp pro leu ala trp asp asn his tyr gly thr
 CTA TTG GAT CCT CTT GCT TGG GAT AAC CAC TAT GGT ACT

                            1580
 gln ile pro lys glu glu trp lys ser gln glu lys ser
 CAG ATA CCA AAA GAA GAG TGG AAA TCC CAA GAG AAG TCA

             1590
 pro glu lys thr ala phe lys lys lys asp thr ile leu
 CCA GAA AAA ACA GCT TTT AAG AAA AAG GAT ACC ATT TTG

 1600                                        1610
 ser leu asn ala cys glu ser asn his ala ile ala ala
 TCC CTG AAC GCT TGT GAA AGC AAT CAT GCA ATA GCA GCA

                                    1620
 ile asn glu gly gln asn lys pro glu ile glu val thr
 ATA AAT GAG GGA CAA AAT AAG CCC GAA ATA GAA GTC ACC
```

FIG. 7(cont'd)

```
                    1630
trp ala lys gln gly arg thr glu arg leu cys ser gln
TGG GCA AAG CAA GGT AGG ACT GAA AGG CTG TGC TCT CAA

    1640                                        1650
asn pro pro val leu lys arg his gln arg glu ile thr
AAC CCA CCA GTC TTG AAA CGC CAT CAA CGG GAA ATA ACT

                                        1660
arg thr thr leu gln ser asp gln glu glu ile asp tyr
CGT ACT ACT CTT CAG TCA GAT CAA GAG GAA ATT GAC TAT

                    1670
asp asp thr ile ser val glu met lys lys glu asp phe
GAT GAT ACC ATA TCA GTT GAA ATG AAG AAG GAA GAT TTT

        1680                                        1690
asp ile tyr asp glu asp glu asn gln ser pro arg ser
GAC ATT TAT GAT GAG GAT GAA AAT CAG AGC CCC CGC AGC

                                        1700
phe gln lys lys thr arg his tyr phe ile ala ala val
TTT CAA AAG AAA ACA CGA CAC TAT TTT ATT GCT GCA GTG

                        1710
glu arg leu trp asp tyr gly met ser ser ser pro his
GAG AGG CTC TGG GAT TAT GGG ATG AGT AGC TCC CCA CAT

            1720
val leu arg asn arg ala gln ser gly ser val pro gln
GTT CTA AGA AAC AGG GCT CAG AGT GGC AGT GTC CCT CAG

1730                                        1740
phe lys lys val val phe gln glu phe thr asp gly ser
TTC AAG AAA GTT GTT TTC CAG GAA TTT ACT GAT GGC TCC
```

FIG. 7(cont'd)

```
                                    1750
phe thr gln pro leu tyr arg gly glu leu asn glu his
TTT ACT CAG CCC TTA TAC CGT GGA GAA CTA AAT GAA CAT

                  1760
leu gly leu leu gly pro tyr ile arg ala glu val glu
TTG GGA CTC CTG GGG CCA TAT ATA AGA GCA GAA GTT GAA

    1770                                        1780
asp asn ile met val thr phe arg asn gln ala ser arg
GAT AAT ATC ATG GTA ACT TTC AGA AAT CAG GCC TCT CGT

                                    1790
pro tyr ser phe tyr ser ser leu ile ser tyr glu glu
CCC TAT TCC TTC TAT TCT AGC CTT ATT TCT TAT GAG GAA

                        1800
asp gln arg gln gly ala glu pro arg lys asn phe val
GAT CAG AGG CAA GGA GCA GAA CCT AGA AAA AAC TTT GTC

        1810                                    1820
lys pro asn glu thr lys thr tyr phe trp lys val gln
AAG CCT AAT GAA ACC AAA ACT TAC TTT TGG AAA GTG CAA

                                        1830
his his met ala pro thr lys asp glu phe asp cys lys
CAT CAT ATG GCA CCC ACT AAA GAT GAG TTT GAC TGC AAA

                            1840
ala trp ala tyr phe ser asp val asp leu glu lys asp
GCC TGG GCT TAT TTC TCT GAT GTT GAC CTG GAA AAA GAT

                1850
val his ser gly leu ile gly pro leu leu val cys his
GTG CAC TCA GGC CTG ATT GGA CCC CTT CTG GTC TGC CAC
```

FIG. 7(cont'd)

```
1860                                                1870
thr asn thr leu asn pro ala his gly arg gln val thr
ACT AAC ACA CTG AAC CCT GCT CAT GGG AGA CAA GTG ACA

                                1880
val gln glu phe ala leu phe phe thr ile phe asp glu
GTA CAG GAA TTT GCT CTG TTT TTC ACC ATC TTT GAT GAG

                    1890
thr lys ser trp tyr phe thr glu asn met glu arg asn
ACC AAA AGC TGG TAC TTC ACT GAA AAT ATG GAA AGA AAC

    1900                                            1910
cys arg ala pro cys asn ile gln met glu asp pro thr
TGC AGG GCT CCC TGC AAT ATC CAG ATG GAA GAT CCC ACT

                                    1920
phe lys glu asn tyr arg phe his ala ile asn gly tyr
TTT AAA GAG AAT TAT CGC TTC CAT GCA ATC AAT GGC TAC

                        1930
ile met asp thr leu pro gly leu val met ala gln asp
ATA ATG GAT ACA CTA CCT GGC TTA GTA ATG GCT CAG GAT

        1940                                        1950
gln arg ile arg trp tyr leu leu ser met gly ser asn
CAA AGG ATT CGA TGG TAT CTG CTC AGC ATG GGC AGC AAT

                                        1960
glu asn ile his ser ile his phe ser gly his val phe
GAA AAC ATC CAT TCT ATT CAT TTC AGT GGA CAT GTG TTC

                        1970
thr val arg lys lys glu glu tyr lys met ala leu tyr
ACT GTA CGA AAA AAA GAG GAG TAT AAA ATG GCA CTG TAC
```

FIG. 7(cont'd)

1980
asn leu tyr pro gly val phe glu thr val glu met leu
AAT CTC TAT CCA GGT GTT TTT GAG ACA GTG GAA ATG TTA

1990 2000
pro ser lys ala gly ile trp arg val glu cys leu ile
CCA TCC AAA GCT GGA ATT TGG CGG GTG GAA TGC CTT ATT

2010
gly glu his leu his ala gly met ser thr leu phe leu
GGC GAG CAT CTA CAT GCT GGG ATG AGC ACA CTT TTT CTG

2020
val tyr ser asn lys cys gln thr pro leu gly met ala
GTG TAC AGC AAT AAG TGT CAG ACT CCC CTG GGA ATG GCT

2030 2040
ser gly his ile arg asp phe gln ile thr ala ser gly
TCT GGA CAC ATT AGA GAT TTT CAG ATT ACA GCT TCA GGA

2050
gln tyr gly gln trp ala pro lys leu ala arg leu his
CAA TAT GGA CAG TGG GCC CCA AAG CTG GCC AGA CTT CAT

2060
tyr ser gly ser ile asn ala trp ser thr lys glu pro
TAT TCC GGA TCA ATC AAT GCC TGG AGC ACC AAG GAG CCC

2070 2080
phe ser trp ile lys val asp leu leu ala pro met ile
TTT TCT TGG ATC AAG GTG GAT CTG TTG GCA CCA ATG ATT

2090
ile his gly ile lys thr gln gly ala arg gln lys phe
ATT CAC GGC ATC AAG ACC CAG GGT GCC CGT CAG AAG TTC

FIG. 7(cont'd)

```
                                2100
ser ser leu tyr ile ser gln phe ile ile met tyr ser
TCC AGC CTC TAC ATC TCT CAG TTT ATC ATC ATG TAT AGT

          2110
leu asp gly lys lys trp gln thr tyr arg gly asn ser
CTT GAT GGG AAG AAG TGG CAG ACT TAT CGA GGA AAT TCC

2120                                          2130
thr gly thr leu met val phe phe gly asn val asp ser
ACT GGA ACC TTA ATG GTC TTC TTT GGC AAT GTG GAT TCA

                                   2140
ser gly ile lys his asn ile phe asn pro pro ile ile
TCT GGG ATA AAA CAC AAT ATT TTT AAC CCT CCA ATT ATT

                    2150
ala arg tyr ile arg leu his pro thr his tyr ser ile
GCT CGA TAC ATC CGT TTG CAC CCA ACT CAT TAT AGC ATT

    2160                                          2170
arg ser thr leu arg met glu leu met gly cys asp leu
CGC AGC ACT CTT CGC ATG GAG TTG ATG GGC TGT GAT TTA

                sphI                    2180
asn ser cys ser met pro leu gly met glu ser lys ala
AAT AGT TGC AGC ATG CCA TTG GGA ATG GAG AGT AAA GCA

                         2190
ile ser asp ala gln ile thr ala ser ser tyr phe thr
ATA TCA GAT GCA CAG ATT ACT GCT TCA TCC TAC TTT ACC

        2200                                      2210
asn met phe ala thr trp ser pro ser lys ala arg leu
AAT ATG TTT GCC ACC TGG TCT CCT TCA AAA GCT CGA CTT
```

FIG. 7(cont'd)

```
                                                2220
his leu gln gly arg ser asn ala trp arg pro gln val
CAC CTC CAA GGG AGG AGT AAT GCC TGG AGA CCT CAG GTG

                              2230
asn asn pro lys glu trp leu gln val asp phe gln lys
AAT AAT CCA AAA GAG TGG CTG CAA GTG GAC TTC CAG AAG

                  2240
thr met lys val thr gly val thr thr gln gly val lys
ACA ATG AAA GTC ACA GGA GTA ACT ACT CAG GGA GTA AAA

2250                                            2260
ser leu leu thr ser met tyr val lys glu phe leu ile
TCT CTG CTT ACC AGC ATG TAT GTG AAG GAG TTC CTC ATC

                                    2270
ser ser ser gln asp gly his gln trp thr leu phe phe
TCC AGC AGT CAA GAT GGC CAT CAG TGG ACT CTC TTT TTT

                       2280
gln asn gly lys val lys val phe gln gly asn gln asp
CAG AAT GGC AAA GTA AAG GTT TTT CAG GGA AAT CAA GAC

    2290                                            2300
ser phe thr pro val val asn ser leu asp pro pro leu
TCC TTC ACA CCT GTG GTG AAC TCT CTA GAC CCA CCG TTA

                          ecoRI      2310
leu thr arg tyr leu arg ile his pro gln ser trp val
CTG ACT CGC TAC CTT CGA ATT CAC CCC CAG AGT TGG GTG

                             2320
his gln ile ala leu arg met glu val leu gly cys glu
CAC CAG ATT GCC CTG AGG ATG GAG GTT CTG GGC TGC GAG

         2330    2332
ala gln asp leu tyr OP
GCA CAG GAC CTC TAC TGA
```

FIG. 7(cont'd)

INTERNATIONAL SEARCH REPORT

International Application No PCT/US87/01814

I. CLASSIFICATION OF SUBJECT MATTER (If several classification symbols apply, indicate all) [3]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(4): A61K 35/14,C12P 21/00, C12P 21/02, C12N 15/00

II. FIELDS SEARCHED

Minimum Documentation Searched [4]

Classification System [5]	Classification Symbols
U.S.	514/2 424/101 530/383 435/68,70,172.3,240,241,253,255,317.1

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [6]

Computer Search CAS, BIOSIS, APS, BIONET, under: Sequence, Factor VIII, deletion, mutat!, clon!, variant, mutein, modified DNA, genetic(w) engineering

III. DOCUMENTS CONSIDERED TO BE RELEVANT [14]

Category * [14]	Citation of Document, [16] with Indication, where appropriate, of the relevant passages [17]	Relevant to Claim No. [18]
Y,P	R.L. Burke et al, "The functional domains of coagulation factor VIII: C," J. Biological Chemistry, Vol. 261 No. 27, pages 12574-12578, published Sept. 25, 1986 by American Society Biological Chemists Inc. (Washington, D.C., USA). See the entire document.	1-19
Y,P	J.J. Toole et al, "A large region (≃95kDa) of human factor VIII is dispensable for in vitro procoagulant activity" Proc. Natl. Acad. Sci. Vol. 83, pages 5939-5942 published August 1986 by National Academy of Science, (Washington, D.C. USA). See the entire document.	1-19

* Special categories of cited documents: [15]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search [2]	Date of Mailing of this International Search Report [2]
19 OCTOBER 1987	03 NOV 1987
International Searching Authority [1]	Signature of Authorized Officer [20]
ISA/US	Robin Lyn Teskin

Form PCT/ISA/210 (second sheet) (October 1981)

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, ¹⁶ with indication, where appropriate, of the relevant passages ¹⁷	Relevant to Claim No ¹⁸
Y,P	WO 86/06101 Published October 23, 1986, Genetics Institute Inc., see the entire document.	1-19
A	J. Gitschier et al, "Characterization of the human factor VIII gene", Nature, Vol. 312, pages 326-330, 22 November 1984, MacMillan Journals LTD (London, England) see the entire document.	1-19
Y	W.I. Wood et al, "Expression of active human factor VIII from recombinant DNA clones" Nature Vol. 312, pages 330-336, 22 November 1984, MacMillan Journals LTD (London, England). See the entire document.	1-19
Y	G.A. Vehar et al, "Structure of human factor VIII", Nature Vol. 312 pages 337-342, 22 November 1984, MacMillan Journals LTD (London, England). See the entire document.	1-19
Y	J.J. Toole et al, "Molecular cloning of a cDNA encoding human anti-haeomophilic factor", Nature, Vol. 312, pages 342-347, 22 November 1984, MacMillan Journals LTD (London, England). See the entire document.	1-19

FURTHER INFORMATION CONTINUED FROM THE SEC ND SHEET

A	D. Eaton et al, "Proteolytic processing of human factor VIII Correlation of specific cleavages with Thrombin, Factor XA, and Activator Protein C With Activation and Inactivation of Factor VIII Coagulant Activity", Biochemicistry, Vol. 25 pages 505-512, published 1986, American Chemical Society (Washington, D.C. USA). See pages 507-512 in particular.	1-19

V. ☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [10]

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. ☐ Claim numbers . because they relate to subject matter [12] not required to be searched by this Authority, namely:

2. ☐ Claim numbers ..... , because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out [13], specifically:

VI. ☐ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [11]

This International Searching Authority found multiple inventions in this international application as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

☐ The additional search fees were accompanied by applicant's protest.

☐ No protest accompanied the payment of additional search fees.